OPTICA ACTA

Vol. 5. Nos 1-2 March-June 1958

Quality evaluations of optical systems

E. H. Linfoot

The Observatories, University of Cambridge, England.

Summary. — *A general theory of optical image evaluation seems now to be within reach of investigation as a result of the introduction into this field of the powerful methods of* Fourier *analysis. The most important practical application of such a theory is to the quality evaluation of optical systems. To be realistic, such a quality evaluation must be defined relative to the statistical properties of a prescribed object set and a prescribed receiving surface. A simplification results from the fact that, in the majority of applications, the quality of a high-grade optical system can be satisfactorily assessed by its power to reproduce low-contrast fine detail. This leads to the idea of assessing the system by means of its averaged performance on low contrast random object sets occupying isoplanatism patches in different parts of the field.*

As is well-known, fine detail is in general reproduced in a metamorphosed form, which however can usually be interpreted correctly by a practised operator. Two distinct types of quality evaluation result according as we demand that the image shall resemble the object as closely as possible, or that it shall give the maximum amount of information about the object without regard to the amount of work (image interpretation) which may be needed to extract this information. In the second type of evaluation, an essentially unique evaluation of image quality exists, viz. that based on the mean information content of the images of a prescribed low-contrast object set ; in the first, there is no unique evaluation, but a number of different evaluations of image quality can be made, each of which leads to a corresponding quality evaluation for the optical system. An exploratory study of the relative merits of these evaluations, and of thier interrelations, seems to be an urgent task for the theoretical optician.

A beginning to this study is attempted in the present paper. Three simple and natural assessments of the first type are considered, and it is shown that in small-field systems the third assessment (correlation quality) is related in an interesting way to a generalised form of Strehl *definition. For such systems, the first assessment (relative structural content) is known to agree with assessment by information content in the limiting case of a random object set whose images are almost completely smothered in random noise. The second assessment (fidelity) is connected with the other two by a simple mathematical relation. The discussion includes consideration of the effects of image spread and noise in the receiving surface.*

Sommaire. — *Il paraît maintenant possible de commencer à bâtir une théorie générale de la qualité des images optiques, du fait de l'introduction dans ce domaine des méthodes puissantes de l'analyse de* Fourier.

L'application pratique la plus importante d'une telle théorie est la détermination de la qualité des systèmes optiques. Pour être réaliste, une telle évaluation de qualité doit être définie à partir de propriétés statistiques d'ensembles d'objets et des propriétés d'une surface réceptrice donnée. Une simplification résulte du fait que, dans la majorité des applications, la qualité des bons systèmes optiques peut être déterminée de façon satisfaisante par leur aptitude à reproduire de petits détails de faible contraste. Ceci conduit à l'idée d'évaluer la qualité d'un système à l'aide de la moyenne des résultats obtenus sur des objets aléatoires de faible contraste occupant des régions isoplanétiques dans différentes parties du champ. Il est bien connu que les fins détails subissent généralement une transformation, mais qu'un observateur entraîné peut souvent les interpréter correctement. Deux types de méthodes d'estimation de la qualité en résultent, suivant que l'on demande que l'image ressemble à l'objet d'aussi près que possible, ou que l'on obtienne le maximum d'information sur l'objet, sans tenir compte de l'importance du travail d'interprétation qui peut être nécessaire pour extraire cette information. Dans le deuxième cas une évaluation unique existe : celle qui est basée sur le calcul de la quantité moyenne d'information dans les images d'un ensemble donné d'objets de faible contraste. Dans le premier cas, il n y a pas d'évaluation unique, mais un certain nombre de méthodes peuvent être employées, chacune conduisant à une évaluation de la qualité du système optique. Une étude des mérites relatifs de ces critères et de leurs relations paraît s'imposer au théoricien.

On présente dans l'article ci-dessous, le début d'une telle étude. Trois critères naturels et simples (appartenant à la première catégorie) sont envisagés et l'on montre que dans les systèmes à faible champ le 3e critère (critère de corrélation) est relié de façon intéressante à une forme généralisée du critère de Strehl. *Pour de tels systèmes, on sait que le premier critère (évaluation du « contenu structural ») s'accorde avec celui qui résulte du calcul de l'information dans le cas limite d'un ensemble d'objets aléatoires dont les images sont presque complètement masquées par le bruit. Le second critère (fidélité) est relié aux deux autres par une relation mathématique simple. La discussion tient compte des effets de la diffusion de l'image et du bruit introduits par la surface réceptrice.*

Zusammenfassung. — *Dank der leistungsfähigen Methode der* Fourier *analyse scheint es jetzt möglich zu sein, eine allgemeine Theorie für die Beurteilung der Bildgüte aufzustellen. Die wichtigste praktische Anwendung einer solchen Theorie richtet sich auf die Leistungsbewertung optischer Systeme. Um in einem realen Rahmen zu bleiben, muss sich diese Leistungsbewertung auf die statistischen Eigenschaften einer vorgegebenen Objektmenge und einer festgelegten Auffangfläche beziehen. Eine Vereinfachung ergibt sich aus dem Umstand, dass bei der Mehrzahl der Anwendungen die Leistung eines wohlkorrigierten optischen Systems hinreichend gut durch seine Wiedergabe kontrastarmer Einzelheiten beurteilt werden kann. Das führt auf den Gedanken, ein optisches System nach der durchschnittlichen Wiedergabe zufälliger kontrastarmer Feinstrukturen in den verschiedenen Isoplanasiebereichen des Gesichtsfeldes zu beurteilen.*

Bekanntlich wird feines Detail im allgemeinen in einer veränderten Form wiedergegeben, die jedoch gewöhnlich von einem erfahrenen Beobachter richtig gedeutet werden kann. Es ergeben sich zwei verschiedene Arten der Gütebeurteilung eines Bildes, je nachdem ob man verlangt, entweder, dass das Bild dem Objekt möglichst ähnlich ist, oder, dass es den grössten Informationsgehalt berü das Objekt liefert, ohne Rücksicht auf den dabei erforderlichen Arbeitsaufwand (Bilddeutung). Bei der zweiten Art der Qualitätsabschätzung gibt es eine im wesentlichen eindeutige Bewertung der Bildgüte, nämlich diejenige, die auf dem mittleren Informationsinhalt der Bilder einer statistisch vorgeschriebenen Menge kontrastarmer Objekte beruht. Die erste Art jedoch bringt keine eindeutige Bewertung, sondern man kann eine Anzahl verschiedener Bildgütekriterien aufstellen, von denen jede zu einer entsprechenden Gütebewertung des optischen Systems führt. Die Untersuchung der jeweiligen Vorzüge dieser Bewertungen und ihrer gegenseitigen Beziehungen scheint für den theoretischen Optiker eine dringende Aufgabe zu sein.

Einen Anfang dazu soll die vorliegende Arbeit geben. Es werden drei einfache und naheliegenden Gütekriterien der ersten Art betrachtet. Es zeigt sich, dass das dritte davon (Korrelationsgüte) bei Systemen mit kleinem Gesichtsfeld in eigenartiger Weise mit einer verallgemeinerten Form der Strehlschen Definitionshelligkeit in Verbindung steht. Für solche Systeme ist bekanntlich das erste Kriterium (relativer Strukturinhalt) im Grenzfall in Übereinstimmung mit der Beurteilung nach dem Informationsinhalt, wenn man sich auf zufällige Objektmengen beschränkt, deren Bilder fast vollständig in den statistischen Schwankungen (noise) untergehen. Das zweite Kriterium (Objekttreue) steht zu den beiden anderen in einer einfachen mathematischen Beziehung. Schliesslich wird noch der Einfluss der Bildverbreiterung und der statistischen Schwankungen in der Auffangfläche betrachtet.

1. **Introduction.** — A general theory of optical image evaluation seems now to be within reach of investigation as a result of the introduction into this field of the powerful methods of FOURIER analysis, which bring into a central position the concept of the spatial frequency response factor of an optical system. This response factor τ is essentially the normalised two-dimensional FOURIER transform of the effective intensity distribution in the diffraction image of a point object by the system [see equation (2.4)]; it is complex-valued in systems which possess unsymmetrical aberrations, and its absolute value $|\tau|$ is the well-known contrast-transmission factor, discussed by COLTMAN (1954) and LINDBERG (1954), among others (1). If there is no asymmetry in the aberrations, nor in the form of the clear aperture, the response factor τ is essentially real-valued (2), but not in general essentially positive. For example, in a heavily defocussed but otherwise aberration-free system with circular aperture, the response factor τ alternates in sign as the spatial frequency increases (HOPKINS, 1955, p. 99) and this gives rise to successive contrast-reversals in the image of a suitable test object. The same is true of defocussed systems with small amounts of spherical aberration (BLACK and LINFOOT, 1957, figs. 3 a, 4 a, 5). Striking photographs illustrating the effect have been published by LINDBERG (1954; fig. 10) and by WASHER (1954; fig. 15.7, 15.8).

The advantages of characterising optical systems by means of their spatial frequency response factors have been outlined in a previous paper (LINFOOT, 1956). It was shown there how this procedure simplifies the analytical problems involved in taking account, in the quality assessment of an optical design, of the statistical properties of the object set on which it is intended to be used and of the properties of the light-sensitive surface which is to receive the images. Quality assessments which do not take account of these can hardly be sufficiently realistic for present day requirements.

Two essentially different types of quality assessment come up for consideration according as we demand that the image shall resemble the object as closely as possible or that it shall contain the maximum amount of information about the object. Different assessments are not, in general, equivalent and it is therefore important to use one which matches the application in view. In the second type of assessment, we regard the systems as a communication channel and assess it according to the mean information content of the images of the statistically prescribed object set (3). This amounts to assessing the system by its *discriminating power*; that is, by the extent to which it can produce distinguishably different images when presented with different members of the object set, without regard to the amount of labour which may be involved in interpreting or "decoding" these images. Information theory here provides an essentially unique assessment of the optical quality of the system.

In the first type of assessment, we ask that the image shall resemble the object as closely as possible. We define the quality of an optical system relative to a prescribed object set as the statistical mean of the quality assessments of the images of the members of the set. To assess in this way the quality of an optical system, we must first know how to assess the quality of the image of an extended object. As already remarked, the choice of assessment needs to be matched to the application in view, and a general exploration of useful assessments and their interrelations appears to be an urgent task for the theoretical optician. Three assessments seem simple and natural enough to deserve examination forthwith; in broad outline they are as follows. Let $\sigma(x, y)$ denote the effective intensity distribution of an extended object and $I(x, y)$ the corresponding intensity distribution in its diffraction image by the optical system. We suppose $I(x, y)$ normalised so that

$$(1.1) \qquad \iint I(x, y)\, dx\, dy = \iint \sigma(x, y)\, dx\, dy.$$

Let A be an isoplanatism patch of the system, that is to say a part of the field, large compared with the size of an AIRY disc, in which the aberrations remain substantially constant. Then three natural assessments of image quality in A are

$$(1.2) \quad T = \iint_A I^2\, dx\, dy \Big/ \iint_A \sigma^2\, dx\, dy \quad \text{(relative structural content)}$$

$$(1.3) \quad \Phi = 1 - \iint_A (\sigma - I)^2\, dx\, dy \Big/ \iint_A \sigma^2\, dx\, dy \quad \text{(fidelity)},$$

$$(1.4) \quad Q = \iint_A \sigma I\, dx\, dy \Big/ \iint_A \sigma^2\, dx\, dy \quad \text{(correlation quality)}.$$

They are connected by the identical relation

$$(1.5) \qquad Q = \frac{1}{2}(T + \Phi).$$

A brief account of these three assessments, and of their extension to non-isoplanatic images, was given in the paper (LINFOOT, 1956) already referred to. Limitations of space prevented a more detailed discussion. In the present paper we examine the properties of three "structural" evaluations, based on the same ideas as (1.2) — (1.4), of the quality of an optical system imaging a random low-contrast object set, and show that the third is related in an interesting way to a generalised form of Strehl definition. The

(1) LINDBERG calls τ the *complex contrast transmission*. Both authors describe the resolution of a test grating as spurious if there is a coarser grating which is not resolved by the same system. In a recent paper (LINFOOT, 1957) the writer has tried a different approach to the problem of defining the resolution of an optical system.

(2) i. e. real-valued for a suitable matching of image surface against object surface. For the effects on τ of mismatch, see section 3 below.

(3) See BLANC-LAPIERRE (1953), FELLGETT and LINFOOT (1955), LINFOOT (1955), BLACK and LINFOOT (1957).

effects of image-spread and noise in the receiving surface are also considered.

2. **Mathematical formalism.** — The notation and basic formulae are substantially the same as those of the paper last mentioned. For convenience, they are given again here, with a few improvements of detail.

A source of negligible size, situated at a point $P = (x, y)$ on the object surface S (which may be at infinity) and emitting unit light-flux of wavelength λ, gives rise to a diffraction image represented by an intensity distribution function $w(x', y' ; x, y ; \lambda)$ over the points (x', y') of the prescribed image surface S'. More precisely, $w(x', y' ; x, y ; \lambda)$ is defined as $1/(c\ dx'\ dy'\ d\lambda)$ times the energy flux from P, in the wave-length range $(\lambda, \lambda + d\lambda)$, reaching, *via* the system, an element $dx'\ dy'$ of area situated at (x', y') in S' where the normalising factor c is chosen so that

$$(2.1) \qquad \int\int_{-\infty}^{\infty} w(x', y' ; x, y ; \lambda)\ dx'\ dy' = 1 .$$

The coordinate mesh-systems (x, y) and (x', y'), in the object and image surfaces respectively, can be defined in a variety of ways. It is convenient to define the coordinates (x, y) of a point P in the object surface as the angular off-axis displacement components (direction cosines) of P as seen from the centre of the GAUSS entry pupil. The wavelength in which the so-called monochromatic aberrations of the system are calculated is called the *principal* wavelength λ_0. A polychromatic pencil of rays issuing from P and traced through the system marks out a polychromatic image patch on the prescribed receiving surface, called the *image patch of P*. The *principal ray* of the pencil is defined as the λ_0-ray which, in the image space, passes through the centre of the exit pupil. Its intersection with the receiving surface will be called the *principal point* P' of the polychromatic image-patch (See fig. 1). It is convenient to define the coordinates x', y' in the receiving surface by assigning to the principal point of the image patch of $P = (x, y)$ the coordinate numbers $x = x'$, $y = y'$. Near the optic axis, the linear distance between two neighbouring points (x_1', y_1'), (x_2', y_2') on the receiving surface is $f\ [(x_1' - x_2')^2 + (y_1' - y_2')^2\ ]^{1/2}$, where f is the focal length of the system. At a point P' in the outer part of the image field, the metric of the (x', y') coordinate system is readily determined to a sufficient approximation by tracing rays from the corresponding object point P and from two nearby object points.

In figure 1, O'' denotes the centre of the GAUSS exit pupil E', O' the intersection of the receiving surface S' with the optic axis O'O'', and O''P' the principal ray of the converging pencil of rays which form the geometrical image patch of the object point $P = (x, y)$ on S (not shown).

With centre P' and radius O'O'' a spherical surface is drawn. The part of this surface, approximately covering the exit pupil, which is intersected by geometrical rays originating from P and passing through the system is called the *spherical reference surface* $M_{P'}$.

Careful selection of coordinate mesh-systems is needed in order to obtain a formally simple analysis which remains valid as a useful approximation in systems of " finite " aperture and field ; for example, in an $F/2$ system working over a field of 10 or 15 degrees

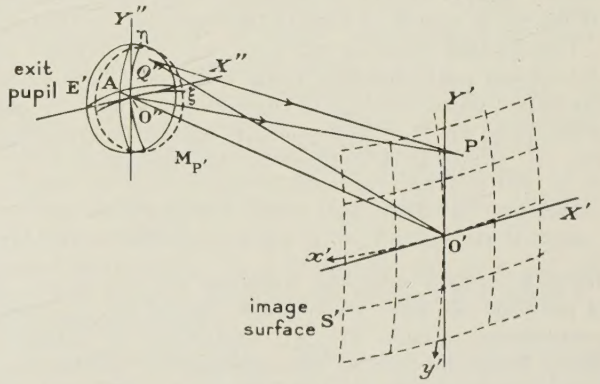

FIG. 1.

diameter. We define a curvilinear-polar coordinate system (θ', ψ') in S' as follows. If $\theta \geqslant 0$ is the off-axis angle of the object-point P as seen from the centre O_1 of the GAUSS entry pupil, and ψ the azimuth (measured from a convenient starting position) of the radial half-plane through the optic axis which contains P and P', then we assign to P' the polar mesh-numbers $\theta' = \theta, \psi' = \psi$. When $\theta = 0$, ψ is indeterminate ; when $\theta > 0$ is small, $x' = f\theta' \cos \psi'$, $y' = f\theta' \sin \psi'$, provided that the ψ-origin is in the half-plane $X''O''O'X'$. When $\theta > 0$ is " finite ", we can write

$$x' = f\theta' \cos \psi' \ [1 + a_2\ \theta'^2 + a_4\ \theta'^4 + \ldots]$$
$$y' = f\theta' \sin \psi' \ [1 + a_2\ \theta'^2 + a_4\ \theta'^4 + \ldots] ,$$

where the constants a_2, a_4, ... depend on S' and on the optical distortion of the system.

We define angular coordinates (ξ, η) of a point Q'' in $M_{P'}$ as follows : Let $\theta'' > 0$ be the measure in radians of the angle Q''P'O'' and ψ'' the angle between the planes P'O''Q'' and P'O''O'. (When Q'' is in the plane P'O''O', on the opposite side of O''P' from O', we take ψ'' to be zero ; when Q'' moves off this plane in the direction of increasing azimuth ψ', we take ψ'' as positive). Then, by definition,

$$\xi = \sin \theta' \cos (\psi' + \psi'')$$
$$\eta = \sin \theta' \sin (\psi' + \psi'').$$

In a system of small field and aperture, (ξ, η) approximate to scale-normalised rectangular Cartesian coordinates in the GAUSS exit pupil, with origin O'' and with O'' ξ, O'' η lying along O''X'', O''Y'' respectively.

The coordinate numbers (ξ, η) in $M_{P'}$ can be used to identify the geometrical rays issuing from P and passing through the system. In light of given wave-

length λ, only one of these rays passes through each point (ξ, η); we call it the (ξ, η)-ray from P. We define the new variables (u, v) by the equations

$$\lambda u = f\xi, \quad \lambda v = f\eta, \tag{2.2}$$

where f is the focal length of the system. To the clear aperture $M_{P'}$ correspond a finite region $\mathfrak{A} = \mathfrak{A}_{x,y;\lambda}$ in the (u, v)-plane. As is well known, (u, v) possess, by virtue of the Huyghens-Kirchhoff theorem, a second interpretation as spatial frequencies relative to the (x', y')-mesh system in the receiving surface [4].

The eikonal function $e(\xi, \eta ; x, y ; \lambda)$ is defined as the optical path distance along a ray of λ-light from the object point $P = (x, y)$ through the system to the point (ξ, η) on $M_{P'}$. The aperture-aberration function $\mathcal{E}(u, v ; x, y ; \lambda)$ is defined by the equation

$$\mathcal{E}(u, v ; x, y ; \lambda) = a(\xi, \eta ; x, y ; \lambda) \exp[-\mathrm{i}ke(\xi, \eta ; x, y ; \lambda)]. \tag{2.3}$$

Here $k = 2\pi/\lambda$ and the function $a(\xi, \eta ; x, y ; \lambda)$ is positive or zero everywhere inside $\mathfrak{A}_{x,y;\lambda}$, zero everywhere outside $\mathfrak{A}_{x,y;\lambda}$. Variations in transparency from one part of the aperture to another can be represented analytically by suitably adjusting the values of the function $a(\xi, \eta ; x, y ; \lambda)$. In optical systems of small aperture and field $a(\xi, \eta ; x, y ; \lambda)$ represents the normalised amplitude of the wave-surfaces at the point (ξ, η) on $M_{P'}$ and is nearly constant when the transparency is nearly uniform. In small-field systems, $\mathfrak{A}_{x,y;\lambda}$ can often be taken to be the circle $u^2 + v^2 = (f^2 \sin^2 \alpha)/\lambda^2$, where $\sin \alpha$ is the numerical aperture.

In light of wavelength λ, the local response factor $\tau(u, v ; x, y ; \lambda)$ is connected with the intensity distribution function $w(x', y', ; x, y ; \lambda)$ by the equation

$$\tau \mathrm{e}^{-2\pi\mathrm{i}(ux+vy)} = \int_{-\infty}^{\infty}\int \mathrm{e}^{-2\pi\mathrm{i}(ux'+vy')} w(x', y' ; x, y ; \lambda)\, \mathrm{d}x'\, \mathrm{d}y' \tag{2.4}$$

and with the aperture-aberration function $\mathcal{E}(u, v ; x, y ; \lambda)$ by the equation

$$\tau = C[\mathcal{E}], \tag{2.5}$$

where $C[\mathcal{E}]$ denotes the (u, v)-autocorrelation function

$$\iint \mathcal{E}(u', v' ; x, y ; \lambda) \mathcal{E}^*(u' + u, v' + v ; x, y ; \lambda)\, \mathrm{d}u'\mathrm{d}v' \tag{2.6}$$

and $\mathcal{E}^*$ denotes the complex conjugate of $\mathcal{E}$. When the system is isoplanatic, the local response factor $\tau(u, v ; x, y ; \lambda)$ reduces to the ordinary response factor $\tau(u, v ; \lambda)$, familiar under a variety of names in the literature [5]. Because $\mathcal{E} = 0$ everywhere outside the finite region $\mathfrak{A}_{x,y;\lambda}$ of the (u, v)-plane, τ is zero at all points of the plane outside the finite region $\mathfrak{T}_{x,y;\lambda}$, defined as the set of points (u, v) such that the intersection $\mathfrak{A}_{x,y;\lambda} \times [\mathfrak{A}_{x,y;\lambda-(u,v)}]$, has a non-zero area (see fig. 2).

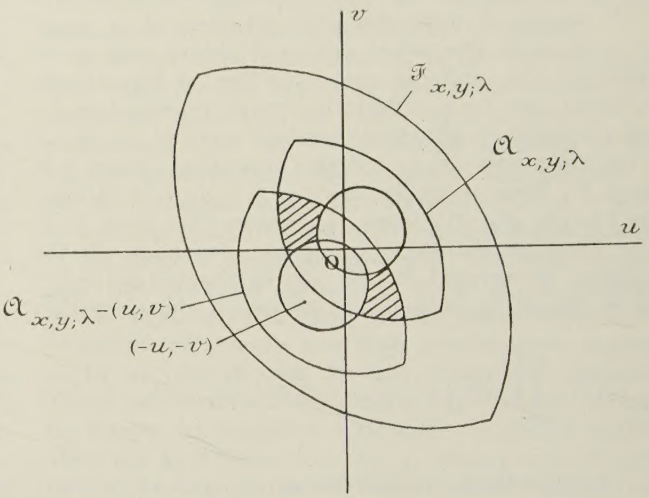

Fig. 2. — The region $\mathfrak{T}_{x,y;\lambda}$ for a system with off-axis edge-vignetting and central obstruction. The intersection $\mathfrak{A}_{x,y;\lambda} \times [\mathfrak{A}_{x,y;\lambda-(u,v)}]$ is indicated by shading.

From its definition, $\mathfrak{T}_{x,y;\lambda}$ is symmetrical about the origin in the (u, v)-plane, whatever the form of the aperture.

Equations (2.3), (2.4) and (2.5) remain valid in systems of "finite" field and aperture provided that the function $a(\xi, \eta ; x, y ; \lambda)$ is suitably reinterpreted as a function, everywhere $\geqslant 0$, which is no longer identical with the normalised amplitude on the wave surface, but which is closely related to this amplitude, is calculable from it in an elementary way, and approximates to it more and more closely as field and aperture are reduced. Its exact form need not concern us here. In many systems of interest to astronomers (for example, in systems of aperture up to $F/2$ and field up to 10 degrees in diameter), the error in taking $a(\xi, \eta ; x, y ; \lambda)$ to represent the normalised amplitude is harmlessly small for present purposes.

Now consider the imaging of a polychromatic object occupying the working field F of an optical system. The object is described by an intensity distribution function $\sigma(x, y ; \lambda)$, which $\geqq 0$ in F and is zero everywhere outside F, and represents, for each value of λ, $1/\mathrm{d}\lambda$ times the flux-density distribution in S of light in the wave-length range $(\lambda, \lambda + \mathrm{d}\lambda)$. The corresponding image-function is given by the equation

$$I(x', y' ; \lambda) = \iint_{\mathrm{F}} \sigma(x, y ; \lambda)\, w(x', y' ; x, y ; \lambda)\, \mathrm{d}x\, \mathrm{d}y. \tag{2.7}$$

[4] The present definitions of ξ, η are preferable to those of my earlier paper because they improve the accuracy of the approximations on which this result depends. The two definitions are equivalent when field and aperture are both very small.

[5] e. g. Duffieux transmission factor, sine-wave response function, complex contrast transmission, contrast transfer factor, frequency response.

For the Fourier transform $F[I]$ of $I(x', y'; \lambda)$ we have the equations

$$F[I] = \int\int_{-\infty}^{\infty} \mathbf{e}^{-2\pi \mathbf{i}(ux+\varphi y)}\, I(x, y; y; \lambda)\, dx\, dy$$

$$= \iint_{F} \sigma(x, y; \lambda)\, dx\, dy \int\int_{-\infty}^{\infty} \mathbf{e}^{-2\pi \mathbf{i}(ux'+\varphi y')} \times w(x', y'; x, y; \lambda)\, dx'\, dy'$$

$$(2.8) = \iint_{F} \sigma(x, y; \lambda)\, \mathbf{e}^{-2\pi \mathbf{i}(ux+\varphi y)}\, \tau(u, \varphi; x, y; \lambda)\, dx\, dy,$$

by (2.4).

In the special case where the object is confined to an isoplanatism patch A of the system, so that σ is zero everywhere outside A, we can replace the domain of integration F by A in (2.8) and can replace $\tau(u, \varphi; x, y; \lambda)$ by $\tau_A(u, \varphi; \lambda)$ in A, obtaining the approximate equation

$$(2.9) \quad F[I] = \tau_A(u, \varphi; \lambda) \iint_{A} \sigma(x, y; \lambda)\, \mathbf{e}^{-2\pi \mathbf{i}(ux+\varphi y)}\, dx\, dy = \tau_A(u, \varphi; \lambda)\, \varepsilon(u, \varphi; \lambda),$$

where

$$(2.10) \quad \varepsilon(u, \varphi; \lambda) = \int\int_{-\infty}^{\infty} \sigma(x, y; \lambda)\, \mathbf{e}^{-2\pi \mathbf{i}(ux+\varphi y)}\, dx\, dy = F[\sigma].$$

is the Fourier transform of $\sigma(x, y; \lambda)$.

Now by Fourier's inversion theorem, (2.10) can be written in the form

$$(2.11) \quad \sigma(x, y; \lambda) = \int\int_{-\infty}^{\infty} \varepsilon(u, \varphi; \lambda)\, \mathbf{e}^{2\pi \mathbf{i}(ux+\varphi y)}\, du\, d\varphi$$

and (2.9) in the form

(2.12)

$$I(x, y; \lambda) = \int\int_{-\infty}^{\infty} \tau_A(u, \varphi; \lambda)\, \varepsilon(u, \varphi; \lambda)\, \mathbf{e}^{2\pi \mathbf{i}(ux+\varphi y)}\, du\, d\varphi.$$

If we interpret the integrals on the right of these two equations as sums, equation (2.11) resolves the object function $\sigma(x, y; \lambda)$ into a sum of Fourier elements $\varepsilon(u, \varphi; \lambda) \exp[2\pi \mathbf{i}(ux + \varphi y)]\, du\, d\varphi$ corresponding to the different spatial frequencies (u, φ), and equation (2.12) shows that each of these elements appears in the image multiplied by the response factor $\tau_A(u, \varphi; \lambda)$. We can describe the general situation by saying that *an optical system acts as a linear filter of spatial frequencies, in which the frequency response factor* $\tau(u, \varphi; x, y; \lambda)$ *varies slowly over the working field* F *and can be treated as invariant (i. e. independent of* x, y*) over each isoplanatism patch* A.

3. **Quality Assessments of Optical Systems.** — It has been pointed out (Linfoot, 1956), that three natural assessments T_F, Φ_F and Q_F of the quality of an optical system relative to a near-monochromatic object set of prescribed statistical mean powerspectrum $\langle |\varepsilon(u, \varphi)|^2 \rangle$ are provided by the $\langle |\varepsilon|^2 \rangle$-weighted (u, φ) means, taken over the whole (u, φ)-plane, of an assessment function $\Psi(\tau)$, where $\tau = \tau(u, \varphi; x, y; \lambda)$ is the local response factor of the system near the field-point (x, y) and λ is the mean wavelength of the near-monochromatic light. The three assessments T_F, Φ_F and Q_F are obtained on giving $\Psi(\tau)$ the three particular forms $|\tau|^2$, $1 - |1 - \tau|^2$ and τ; they correspond respectively to the notions of *relative structural content*, *fidelity*, and *correlation quality* of the image, defined in the same paper (6).

A special case of particular importance is that of a *low-contrast* random object set with a prescribed mean structural power spectrum. Such a set is represented analytically by a statistical set $\{\sigma(x, y)\}$ of functions $\sigma(x, y)$, vanishing outside F and $\geqq 0$ inside F, with the properties :

(1) the statistical mean $\langle \sigma(x, y) \rangle$ has the same value $B_0 > 0$ at all points (x, y) inside F, or, what is equivalent, the statistical average of the mean brightness

$$(3.1) \quad B_{\sigma; D} = \frac{1}{|D|} \iint_{D} \sigma(x, y)\, dx\, dy$$

has the same value B_0 for all regions D interior to F;

(2) the functions which describe object *structure* in terms of spatial frequencies (u, φ), viz. the functions

$$(3.2) \quad \varepsilon'(u, \varphi) = F[\sigma'],$$

where

$$(3.3) \quad \sigma' = \sigma - B_\sigma \varkappa_F, \qquad B_\sigma = \frac{1}{|F|} \iint_{F} \sigma\, dx dy$$

and

$$(3.4) \quad \varkappa_F = \varkappa_F(x, y) = 1 \quad (x, y) \text{ in F}; \quad = 0 \quad (x, y) \text{ not in F},$$

are substantially uncorrelated with B_σ in the sense that for each u, φ

$$(3.5) \quad \langle B_\sigma\, \varepsilon'(u, \varphi) \rangle = 0$$

to a sufficient approximation (7).

(3) the structure-functions $\varepsilon'(u, \varphi)$ satisfy the equation

$$(3.6) \quad \langle |\varepsilon'(u, \varphi)|^2 \rangle \simeq \varphi(u, \varphi),$$

(6) On pp 751-2 of the paper, T_F, Φ_F and Q_F were described in error as the $\langle |\varepsilon|^2 \rangle$ — weighted means of $|\tau|^2$, $1 - |1 - \tau|^2$ and τ over $F \times \mathcal{F}$ instead of over $F \times \mathcal{Q}$, where $\mathcal{Q}$ is the full (u, v) — plane. There is also a mistake in eq. (5.6), where both double integrals should be taken over $\mathcal{Q}$.

(7) This condition is satisfied when there is no tendency for the phase of $\varepsilon'(u, v)$ to be correlated with the mean brightness B_σ.

where the prescribed mean structural power spectrum $\varphi(u, v)$ is a smooth non-negative function such that $\varphi(-u, -v) = \varphi(u, v)$ and

$$(3.7) \qquad \iint_{-\infty}^{\infty} \varphi(u, v)\, du\, dv \ll |\, F\, |\, B_0^2 .$$

In a practical case $\varphi(u, v)$ usually takes the more special form $\varphi(u^2 + v^2) = \varphi(r^2)$. We write $\simeq$ and not $=$ in (3.6) because the fact that each $\sigma'(x, y)$ is confined to F involves the band-limitedness of $\varepsilon'(u, v)$, in consequence of which $\varepsilon'(u, v)$ is uniquely determined by its values at the sampling points

$$(u_p, v_q) = \left(\frac{p}{2a}, \frac{q}{2b}\right),$$

where $p, q = 0, \pm 1, \pm 2, \ldots$ and $2a$, $2b$ are respectively the differences between the greatest and least values of x and of y in F. It seems out of place to go further into these refinements here, beyond remarking that the condition that $\varphi(u, v)$ shall be smooth ensures that they can be ignored without serious inaccuracy.

From (3.2), (3.4) we see that the spectral function $\varepsilon(u, v)$ and the structural density function $\varepsilon'(u, v)$ are connected by the relations [8]

$$(3.8) \qquad \langle\, |\, \varepsilon\, |^2 \,\rangle = \langle\, |\, \varepsilon'\, |^2 \,\rangle + B^2\, |\, F[\, \varkappa_F\,]\, |^2,$$

where

$$(3.9) \qquad B = (\langle\, B_\sigma^2 \,\rangle)^{1/2}.$$

Evidently

$$B^2 = \langle\, B_\sigma^2 \,\rangle = \langle\, (B_\sigma - B_0)^2 \,\rangle + B_0^2 \geqq B_0^2 ;$$

we assume that the low-contrast set $\{\, \sigma \,\}$ satisfies the final condition :

$$(3.10) \qquad \langle\, (B_\sigma - B_0)^2 \,\rangle \ll B_2^0.$$

In ordinary photographic applications this assumption represents an expectation that nearly all the photographs wil be given nearly correct exposures.

Since

$$(3.11) \qquad \iint_{\mathcal{D}} |\, F[\, K_F\,]\, |^2\, du\, dv \simeq \begin{cases} |\, F\, | \\ 0 \end{cases}$$

and hence, by (3.8),

$$(3.12) \qquad \iint_{\mathcal{D}} \langle\, |\, \varepsilon\, |^2 \,\rangle du\, dv \simeq \iint_{\mathcal{D}} \langle\, |\, \varepsilon'\, |^2 \,\rangle du\, dv + \begin{cases} |\, F\, | B^2 \\ 0 \end{cases}$$

according as the point $(u, v) = (0, 0)$ is inside or outside the finite (not small) domain $\mathcal{D}$ of the frequency plane, it follows that for this type of object set

$$T_F,\ \Phi_F,\ Q_F \simeq \frac{\iint_F dx\, dy \iint_{-\infty}^{\infty} \Psi(\tau)\, (\langle\, |\, \varepsilon'\, |^2 \,\rangle + B^2\, |\, F[\, K_F\,]\, |^2)\, du\, dv}{|\, F\, | \iint_{-\infty}^{\infty} (\langle\, |\, \varepsilon'\, |^2 \,\rangle + B^2\, |\, F[\, K_F\,]\, |^2)\, du\, dv},$$

where $\Psi(\tau)$ is given the special forms $|\, \tau\, |^2$, $1 - |\, 1 - \tau\, |^2$, τ respectively in the three assessments,

$$(3.13) \qquad \simeq \frac{B^2 + \iint_{-\infty}^{\infty} \varphi(u, v) \left(|\, F\, |^{-1} \iint_F \Psi(\tau)\, dx\, dy \right) du\, dv}{B^2 + \iint_{-\infty}^{\infty} \varphi(u, v)\, du\, dv}$$

because τ is slowly varying and $\tau(0, 0\, ;\, x, y\, ;\, \lambda) = 1$.

In most practical applications, interest centres round the structure ε' and it is therefore more appropriate to use, in place of T_F, Φ_F and Q_F, the " structural assessments ".

$$(3.14) \qquad T_F',\ \Phi_F',\ Q_F' = \left(\iint_{-\infty}^{\infty} \varphi(u, v)\, \frac{1}{|\, F\, |} \iint_F \Psi(\tau)\, dx\, dy\, du\, dv \right) \Big/ \iint_{-\infty}^{\infty} \varphi(u, v)\, du\, dv ,$$

where as before $\Psi(\tau)$ is given the three special values $|\, \tau\, |^2$, $1 - |\, 1 - \tau\, |^2$, τ respectively in the three assessments. In all three cases, $\Psi(\tau)$ is zero when (u, v) is outside $\mathcal{F}$.

By (3.13) this is equivalent to carrying out the assessments T_F, Φ_F, Q_F on the " reduced " set $\{\, \sigma' \,\}$, where $\sigma' = \sigma - B_\sigma \varkappa_F$, $\varepsilon' = F[\, \sigma'\,]$.

From (3.13) and (3.14) it will be seen that to maximise T_F, Φ_F or Q_F by varying the available parameters of the optical design is substantially equivalent to maximising T_F', Φ_F' or Q_F' respectively ; and that when aperture and field are prescribed these quantities differ from the φ-weighted means over $F \times \mathcal{F}$ of $|\, \tau\, |^2$, $1 - |\, 1 - \tau\, |^2$ or τ, which we shall denote by T_F'', Φ_F'' and Q_F'' respectively, only by the multiplying factor

$$(3.15) \qquad K(\varphi\, ;\, \mathcal{F}) = \iint_{\mathcal{F}} \varphi(u, v)\, du\, dv \Big/ \iint_{-\infty}^{\infty} \varphi(u, v)\, du\, dv ,$$

(8) We have

$$\varepsilon = \varepsilon' + B_\sigma F[\, K_F\,] ;\ |\, \varepsilon\, |^2 = |\, \varepsilon'\, |^2 + B_\sigma^2\, |\, F[\, K_F\,]\, |^2 + 2\, B_\sigma\, \mathcal{R}\{\, \varepsilon'\, F[\, K_F\,]\,\} ;\ \langle\, |\, \varepsilon\, |^2 \,\rangle = \langle\, |\, \varepsilon'\, |^2 \,\rangle + \langle\, B_\sigma^2 \,\rangle\, |\, F[\, K_F\,]\, |^2 + 2\, \mathcal{R}\{\, \langle\, \varepsilon'\, B_\sigma \,\rangle\, F[\, K_F\,]\,\},$$

which yields (3.8) since $\langle\, \varepsilon'\, B_\sigma \,\rangle = 0$ (condition 2).

which is independent of the optical aberrations. It follows that each of the means T''_F, Φ''_F and Q''_F provides a physically meaningful quality assessment for an optical system of prescribed field and aperture, relative to an object set of given spatial power spectrum $\varphi(u, v)$ in $\mathcal{F}$.

A case of particular interest is that in which the optical system has a prescribed aperture and focal length, in which there is no previous knowledge of the statistics of the intended object set beyond the fact that almost all the fine detail is of low contrast, and in which it is this fine detail to which the user attaches importance. Then something near to the best result is obtained by optimising the optical design for use on a low-contrast object set which is of maximum randomness under the constraint of prescribed total structure power in the set of spatial frequencies which can be passed by the system ([9]).

To obtain an analytical representation of such an object set $\{ \sigma \}$ we group its members into subclasses of mutually indistinguishable objects (i. e., objects which all yield the same image with the given system) and use, as the representative of each subclass, the convolution of any one of its members with $F^*(\varkappa_{\mathcal{F}})$, where $\varkappa_{\mathcal{F}} = \varkappa_{\mathcal{F}}(u, v) = 1$ or 0 according as (u, v) is or is not in $\mathcal{F}$. The representative then depends only on the sub class, not on the choice of member. The representative of the sub class to which a particular σ belongs can be written in the form

$$\sigma_{\mathcal{F}} = \sigma * F^* [\varkappa_{\mathcal{F}}],$$

where the lower asterisk denotes a convolution.

The constraint of a prescribed statistical mean power spectrum $\varphi(u, v)$ in $\mathcal{F}$ affects $\{ \sigma \}$ only through the subclass representatives $\sigma_{\mathcal{F}}$, and we interpret the statement that $\{ \sigma \}$ is random apart from this constraint to mean that it satisfies (1), (2) and the further condition that the relative probability in $\{ \sigma \}$ of any subset of σ is equal to the relative probability in $\{\sigma_{\mathcal{F}}\}$ of the corresponding subset of $\sigma_{\mathcal{F}}$. It then follows that the statistical mean over $\{ \sigma \}$ of any expression involving σ only through $\sigma_{\mathcal{F}}$ is equal to the statistical mean over $\{ \sigma_{\mathcal{F}} \}$ of that expression, where $\{ \sigma_{\mathcal{F}} \}$ is random in the sense of maximum entropy under the constraint of given statistical mean power spectrum $\varphi(u, v)$ in the spectral domain $\mathcal{F}$.

The constraint of a prescribed total statistical mean power over $\mathcal{F}$ likewise affects $\{ \sigma \}$ only through the $\sigma_{\mathcal{F}}$, and the randomness of $\{ \sigma \}$ under this constraint is interpreted as before except that now $\{ \sigma_{\mathcal{F}} \}$ is random in the sense of maximum entropy under the constraint of given total statistical mean power in $\mathcal{F}$. It is well-known that in this case an entropy-maximised $\{ \sigma_{\mathcal{F}} \}$ has a " flat " power spectrum in $\mathcal{F}$, i. e. $\varphi(u, v)$ is constant in $\mathcal{F}$ for such a set.

Since T''_F, Φ''_F and Q''_F involve σ only through $\sigma_{\mathcal{F}}$ (because the values of $\langle\, |\, \varepsilon'(u, v)\,|^2 \,\rangle$ in $\mathcal{F}$ depend only on $\sigma_{\mathcal{F}}$) it follows that an analytical representation of such an object set as a special case of a set satisfying (1), (2), (3) and (4) above is obtained, to a sufficient approximation for the present purpose, by supposing (what cannot be strictly true) that $\varphi(u, v)$ is constant in $\mathcal{F}$ and is zero outside $\mathcal{F}$; that is, that the set $\{ \sigma - B_\sigma \varkappa_{\mathcal{F}} \}$ has a uniform statistical mean power spectrum effectively confined to $\mathcal{F}$. On this assumption, we obtain from (3.14) the approximate evaluations

$$(3.16) \quad T''_F = \underset{F,\,\mathcal{F}}{\text{Mean}}\, |\tau|^2$$

$$(3.17) \quad \Phi''_F = \underset{F,\,\mathcal{F}}{\text{Mean}}\, (1 - |1 - \tau|^2)$$

$$(3.18) \quad Q''_F = \underset{F,\,\mathcal{F}}{\text{Mean}}\, \tau = \underset{F,\,\mathcal{F}}{\text{Mean}}\, \mathcal{R}(\tau), \ ([10])$$

where $\tau(u, v ; x, y ; \lambda)$ is the local response factor of the optical system.

The expressions on the right of (3.16), (3.18) assess, in three different ways, the *quality* of an optical system of prescribed field and aperture in terms of its performance on a low-contrast random object set. They are connected by the identical relation

$$T''_F + \Phi''_F - 2\, Q''_F = 0.$$

It will be seen that T''_F is unchanged in value when a small, smooth change is made in the " matching " between S′ and S. For the effect of such a change is to replace τ by $\tau\, \mathbf{e}^{2\pi \mathbf{i}(u\,\delta x + v\,\delta y)}$, where $x' = x + \delta x$, $y' = y + \delta y$ are the old coordinates in S′ of the point which is now matched against (x, y) in S, and the value of $|\tau|^2$ remains unchanged. On the other hand, Φ''_F and $Q''_F = \frac{1}{2}(T''_F + \Phi''_F)$ are changed in value by a change in matching ; and this point is taken up again later.

Since τ vanishes outside $\mathcal{F}$, the right hand side of (3. 16) is $1/|\mathcal{F}|$ times the so-called structural resolution of the optical system (Linfoot, 1957), and therefore agrees, apart from a normalizing factor, with the mean information density of the images of a low-contrast random object set when fine image-detail is almost smothered in random noise.

The right hand side of (3.18) is related in an equally interesting manner to the Strehl definition. A few historical remarks about Strehl definition may be in place at this point.

Suppose that an optical system with circular aperture images a luminous point on to a prescribed receiving surface. If aberrations, including defect of focus,

([9]) The same optimisation gives something near to the best result for the more special type of object set, of particular nterest to astronomers, in which the members consist of faint luminous points randomly distributed over a uniform background.

([10]) Because $\tau(-u, -v ; x, y ; \lambda)$ is equal to the complex conjugate of $\tau(u, v ; x, y ; \lambda)$.

are absent, the image in light of wavelength λ takes the well-known form of a bright central nucleus (the " Airy disc ") which contains 84 % of the light, and a series of faint concentric rings whose radii increase by approximately equal amounts. When the wavelength λ of the light and the numerical aperture $\sin\alpha$ of the system are varied, the Airy diffraction pattern simply changes scale, the diameter of the Airy disc being given by the expression $1.22\,\lambda/\sin\alpha$.

When an extended, self-luminous, polychromatic object is imaged by a system of negligible aberrations, including chromatism, the image may be regarded as built up out of overlapping Airy discs, of different sizes and containing 84 % of the light, together with their ring systems, which contain the remaining 16 %. It follows that fine detail in the object necessarily appears in the image with reduced contrast, and that detail below a certain size is no longer recognizable.

The main effect of a small amount of defocussing, or of other symmetrical aberration, on the image of a bright point is to cause a redistribution of light between the disc and rings of the Airy pattern, without much altering their sizes or relative positions. For example [1], an amount 1/4 λ of defocussing (the Rayleigh tolerance limit) or λ of primary spherical aberration at " mean " focus (wave distortion

$$-\lambda r^2 + \lambda r^4,\ 0 \leqq r < 1)$$

diverts a further 17 % of the light from the monochromatic Airy disc into the rings, without appreciably changing the size of the disc or the relative distribution of intensity inside it. Because of the relative faintness of the diffraction rings (in the aberration-free image their brightness is everywhere less than 2 % of that at the centre of the Airy disc), this redistribution of light hardly affects the resolving power of the system when tested on an object consisting of two luminous points very close together. But an increase from 16 to 33 in the percentage of light diffracted outside the bright nucleus of the images of point objects produces a decided loss of contrast in the fine detail of the image of an extended object. Therefore, as was pointed out by Conrady (l. c., p. 591), the Rayleigh quarter-wave limit can scarcely be regarded as a safe design-tolerance for instruments, such as microscopes, intended to produce high quality images of extended objects.

For such instruments Strehl (1902) proposed a more appropriate criterion of quality, called by him the Definitionshelligkeit and defined as the ratio of the maximum intensity in the actual diffraction image of a point source by the system to the value which this intensity would have if the aberrations were removed. For small aberrations of the type met with in the practical design of high-quality micro-objectives and similar systems, the central nuclei of these two images are of about the same size and enclose approximately the same relative intensity distribution ; it follows that in such systems the Definitionshelligkeit or Strehl *definition* S gives a measure of the fraction of the light contained in the central nucleus of the image. In monochromatic light, this fraction is $0{,}84\,S$, while the quantity $\delta = 1.00 - 0.84\,S$ represents the fraction of the light which is diffracted outside the nucleus. To this order of approximation, the effect of aberrations on the nucleus of each point image is to decrease its brightness in a fixed ratio, the lost light reappearing in the faint outer part of the image, which is altered in a more complicated manner.

Thinking once more of the image as essentially built up out of bright nuclei, we can regard the loss of image-quality which experience shows to result from a reduction of Strehl definition in high-grade systems as due to a loosening of the correlation between the intensity distributions in the object and image surfaces. For, in the extended image, each contributory bright nucleus is overlaid by light from the faint outer parts of a large number of neighbouring elementary images : as S decreases, the amount of this " spread " light increases and the correlation between object and image is thereby made looser. It will now be shown how this statement can be given a more precise meaning, and one which remains valid when the aberrations are no longer restricted to be small.

We first observe that when the wave aberrations $\varphi(\xi, \eta : x, y)$ of a monochromat with negligible or uniform absorption are symmetrical at each point (x, y) of the field and everywhere less than $\frac{1}{2}\lambda$, the intensity of the diffraction image of each luminous object point $P = (x, y)$ on S is brightest at the central point P′ of the image, which is the same as the intersection with S′ of the principal ray of the near-stigmatic converging pencil which forms the geometrical image of P. In this case, we obtain by assigning to P′ the coordinate numbers $x' = x$, $y' = y$ a coordinate mesh-system in S′ which matches each object point against the brightest point in its diffraction image.

The Strehl definition is here

$$\left|\iint_{\mathfrak{A}} \exp\left\{2\pi \mathbf{i}\,\varphi(\xi\,\eta\,;\,x,\,y)\right\} du\,d\upsilon\right|^2 \Big/ \left(\iint_{\mathfrak{A}} du\,d\upsilon\right)^2$$

$$= \left|\frac{1}{|\mathfrak{A}|}\iint_{\mathfrak{A}} \exp\left\{2\pi \mathbf{i}\,\varphi(\xi\,\eta\,;\,x,\,y)\right\} du\,d\upsilon\right|^2. \tag{3.19}$$

Now from equation (2.5), namely

$$\tau = \int_{-\infty}^{\infty}\!\!\int \mathfrak{E}(u', \upsilon'\,;\,x, y\,;\,\lambda) \times$$

$$\mathfrak{E}^*\,(u' + u, \upsilon' + \upsilon\,;\,x, y\,;\,\lambda)\,du'\,d\upsilon',$$

[1] A. E. Conrady, *R. A. S. Monthly Notices*, **79**, 1919, 575.

we have in general

$$(3.20)\qquad \iint_{\mathcal{F}} \tau \, du \, d\nu = \int_{-\infty}^{\infty}\int \tau \, du \, d\nu = \left| \int_{-\infty}^{\infty}\int_{\mathcal{A}} \mathcal{E}(u, \nu ; x, y ; \lambda) \, du \, d\nu \right|^2 ;$$

in the present case this becomes

$$(3.21)\qquad \iint_{\mathcal{F}} \tau \, du \, d\nu = \left| \iint_{\mathcal{A}} \exp\left\{2\pi \mathbf{i} \, \varphi(\xi, \eta ; x, y)\right\} du \, d\nu \right|^2 \Big/ \iint_{\mathcal{A}} du \, d\nu.$$

(The denominator takes account of the normalisation adopted for $\mathcal{E}$). From (3.19) and (3.21) we see that, with the adopted matching of S′ against S, the value of the integral

$$\iint_{\mathcal{F}} \tau \, du \, d\nu$$

at each point (x, y) of the field is equal to $|\mathcal{A}|$ times the STREHL definition at that point.

This result is capable of a wide extension. Suppose for example that the wave-aberrations of the system are no longer restricted to be symmetrical, or to be small; that chromatism is present; and that the transmission is not restricted to be uniform, so that edge-vignetting, central obstructions and aperture shading are also allowed. Suppose further that we are interested in assessing the performance of the system in light of a prescribed effective spectral distribution $\rho(\lambda)$ [11].

We can define the *effective* STREHL *definition* of the image of a polychromatic point source P situated at (x, y) in the object surface S as $\widehat{I}(x, y)/\widehat{I}_0(x, y)$, where $\widehat{I}_0(x, y)$ is the maximum effective intensity in the actual diffraction image of P and $\widehat{I}_0(x, y)$ is the maximum effective intensity in the image obtained when all phase-distortions are removed from the wave-surfaces issuing from the exit pupil [12] and only the variations in amplitude are left. Classical STREHL definition is the special case in which the light is monochromatic and the amplitude is constant over the wave surfaces issuing from the exit pupil of the system.

By the *extended* STREHL *definition* of an optical system over the field F, we shall mean the mean value, as (x, y) explores F, of the effective STREHL definition of the image of a point source at (x, y). Evidently the STREHL definition in this extended sense depends on the effective spectral distribution $\rho(\lambda)$ of the light and on the chromatism of the system, as well as on the so-called monochromatic aberrations. When the system is isoplanatic over the field F, the effective STREHL definition relative to a given spectral distribution is the same at each point of F and agrees in value with the extended STREHL definition over F.

Now consider a general centred system with aperture aberration function (2.3). In substantially monochromatic light of wavelength λ, the quotient

$$(3.22)\qquad \left| \int_{-\infty}^{\infty}\int \mathcal{E}(u, \nu ; x, y ; \lambda) \, du \, d\nu \right|^2 \Big/ \left(\int_{-\infty}^{\infty}\int \left| \mathcal{E}(u, \nu ; x, y ; \lambda) \right| du \, d\nu \right)^2,$$

which is always $\leqq 1$, is equal to the ratio

$$(3.23)\qquad w(x', y' ; x, y ; \lambda) / \operatorname*{Max}_{(x', y')} w_0(x', y' ; x, y ; \lambda)$$

and this is equal to the extended STREHL definition $S(\mathrm{P})$ of the image of P if (and only if) the matching of S′ against S is so chosen that the point $x' = x$, $y' = y$ is an absolute maximum of the image-intensity function $w(x', y' ; x, y ; \lambda)$. For reasons which will be clear later, we call a choice of matching which satisfies this condition an *optimal matching* in light of wavelength λ.

The relation (3.20) connecting τ and $\mathcal{E}$, viz.

$$(3.24)\qquad \iint_{\mathcal{F}} \tau \, du \, d\nu = \left| \int_{-\infty}^{\infty}\int \mathcal{E}(u, \nu ; x, y ; \lambda) \, du \, d\nu \right|^2,$$

can be written in the homogeneous form

$$(3.25)\qquad \iint_{\mathcal{F}} \tau \, du \, d\nu = \left| \int_{-\infty}^{\infty}\int \mathcal{E}(u, \nu ; x, y, \lambda) \, du \, d\nu \right|^2 \Big/ \int_{-\infty}^{\infty}\int |\mathcal{E}(u, \nu ; x, y ; \lambda)|^2 \, du \, d\nu,$$

since the normalisation $\tau(0, 0 ; x, y ; \lambda) = 1$, together with (2.5), implies that $\int_{-\infty}^{\infty}\int |\mathcal{E}(u, \nu ; x, y ; \lambda)|^2 \, du \, d\nu = 1$,

which is our normalisation of $\mathcal{E}$. By introducing a further factor $\left(\int_{-\infty}^{\infty}\int |\mathcal{E}(u, \nu ; x, y ; \lambda)| \, du \, d\nu \right)^2$ in nume-

[11] " effective " refers to a possible variation with λ of the response of the sensitive surface S′, which is presumed not to possess direct colour-distrimination.

[12] That is, the wave-fronts become spheres centred at a. of S′ in or very near to the core of the previous diffraction image. The exact position chosen for these centres does not affect the results.

rator and denominator, (3.25) can be given the form

$$\iint_{\mathcal{F}} \tau \, du \, d\varphi = \frac{\left| \iint \mathcal{E} \, du \, d\varphi \right|^2}{\left(\iint |\mathcal{E}| \, du \, d\varphi \right)^2} \cdot \frac{\left(\iint |\mathcal{E}| \, du \, d\varphi \right)^2}{\iint |\mathcal{E}|^2 \, du \, d\varphi}. \tag{3.26}$$

Now suppose in particular that the matching is optimal near (x, y). Then the first factor on the right of (3.26) becomes $S(\mathrm{P})$ and we obtain the equation

$$S(\mathrm{P}) = U_{x,y} \operatorname*{Mean}_{\mathcal{F}} \tau, \tag{3.27}$$

where, by Schwarz's inequality,

$$\begin{aligned} U_{x,y} &= |\mathcal{F}| \iint_{-\infty}^{\infty} |\mathcal{E}|^2 \, du \, d\varphi \Big/ \left(\iint_{-\infty}^{\infty} |\mathcal{E}| \, du \, d\varphi \right)^2 \\ &= |\mathcal{F}| \iint_{\mathfrak{A}} |\mathcal{E}|^2 \, du \, d\varphi \Big/ \left(\iint_{\mathfrak{A}} |\mathcal{E}| \, du \, d\varphi \right)^2 \\ &\geqq |\mathcal{F}| / |\mathfrak{A}|, \end{aligned} \tag{3.28}$$

with equality only if $|\mathcal{E}|$ is constant over $\mathfrak{A}$. When $\mathfrak{A}$ is a circle or a rectangle, $\mathcal{F}$ is a similar area of twice the linear size, $|\mathcal{F}|/|\mathfrak{A}| = 4$ and $U_{x,y} \geqq 4$. Then by (3.27)

$$\operatorname*{Mean}_{\mathcal{F}} \tau \leqq \frac{1}{4} S(\mathrm{P}), \tag{3.29}$$

with equality only if the matching is optimal near (x, y) and $|\mathcal{E}|$ is constant over $\mathfrak{A}$. It follows that the extended Strehl definition

$$S_{\mathrm{F}} = \operatorname*{Mean}_{\mathrm{F}} S(\mathrm{P}) \geqq 4 \operatorname*{Mean}_{\mathrm{F}, \mathcal{F}} \tau, \tag{3.30}$$

with equality only if the matching is optimal over F, and the absorption is substantially uniform over the aperture ([13]). Now by (3.18)

$$Q''_{\mathrm{F}} = \operatorname*{Mean}_{\mathrm{F}, \mathcal{F}} \tau \tag{3.31}$$

for a random object set, whatever the matching. It follows that *for a small-field system used on a monochromatic random object set, the maximised value of* Q''_{F} *is equal to* 1/4 *of the extended* Strehl *definition over* F, *and so never exceeds* 1/4.

The correlation quality of an optical system relative to any object set is defined as the statistical mean of the correlation qualities of the images of the members of the set, and the above arguments can be extended to the practically important case of a system intended for use on a random object set in which the statistical mean effective spectral distribution $\rho(\lambda)$ is the same at all points of the field, and structure and colour are uncorrelated, while the receiving surface does not possess direct colour discrimination. We replace $w(x', y' ; x, y ; \lambda)$, $w_0(x', y' ; x, y ; \lambda)$, $\tau(u, \varphi ; x, y ; \lambda)$, $\tau_0(u, \varphi, x, y, \lambda)$ by their respective $\rho(\lambda)$ -weighted λ-means

$$w(x', y' ; x, y), \; w_0(x', y' ; x, y), \; \tau(u, \varphi ; x, y), \; \tau_0(u, \varphi ; x, y),$$

and define the *effective correlation quality* Q''_{F} of the system relative to the set as the $(\lambda^{-2}\rho(\lambda))$-weighted λ-mean of the expression $\operatorname*{Mean}_{\mathrm{F}, \mathcal{F}} \tau$ in (3.18). This is equivalent to defining the effective correlation quality Q''_{F} relative to the polychromatic set by the equation

$$Q''_{\mathrm{F}} = \operatorname*{Mean}_{\mathrm{F}} \left(\iint_{-\infty}^{\infty} \tau(u, \varphi ; x, y) \, du \, d\varphi \Big/ \operatorname*{Mean}_{\rho(\lambda)} |\mathcal{F}_{x,y;\lambda}| \right), \tag{3.32}$$

in which $\tau(u, \varphi ; x, y)$ is the polychromatic response function (see Linfoot, 1956) and $\operatorname*{Mean}_{\rho(\lambda)}$ stands for the $\rho(\lambda)$-weighted λ-mean.

The *effective structural content* T''_{F} and the *effective fidelity* Φ''_{F} of the system relative to the same object set are obtained on replacing τ in (3.32) by $|\tau|^2$ and $1 - |1 - \tau|^2$ respectively. Evidently the relation

$$T''_{\mathrm{F}} + \Phi''_{\mathrm{F}} - 2 Q''_{\mathrm{F}} = 0 \tag{3.33}$$

is still satisfied, and the generalised T''_{F} is still invariant under a small, smooth change of matching.

Now it is easy to verify that $w_0(x, y ; x, y ; \lambda) = \lambda^{-2} H(x, y)$, where the positive function $H(x, y)$ is independent of λ. Therefore

$$\begin{aligned} \frac{w(x, y ; x, y)}{w_0(x, y ; x, y)} &= \frac{\int \rho(\lambda) \, w(x, y ; x, y ; \lambda) \, d\lambda}{\int \rho(\lambda) \, w_0(x, y ; x, y ; \lambda) \, d\lambda} \\ &= \frac{1}{H(x, y)} \frac{\int \rho(\lambda) \, w(x, y ; x, y ; \lambda) \, d\lambda}{\int \lambda^{-2} \, \rho(\lambda) \, d\lambda} \\ &= \int \lambda^{-2} \, \rho(\lambda) \, \frac{w(x, y ; x, y ; \lambda)}{w_0(x, y ; x, y ; \lambda)} \, d\lambda \Big/ \int \lambda^{-2} \, \rho(\lambda) \, d\lambda \\ &= \operatorname*{Mean}_{\lambda^{-2}\rho(\lambda)} \frac{w(x, y ; x, y ; \lambda)}{w_0(x, y ; x, y ; \lambda)}, \end{aligned}$$

([13]) In optical systems of focal ratio shorter than $F/2$, the constancy of $|\mathcal{E}|$ over $\mathfrak{A}$ has a less simple meaning.

where $\underset{\lambda^{-2}\rho(\lambda)}{\text{Mean}}$ denotes the $(\lambda^{-2}\rho(\lambda))$ -weighted λ-mean,

$$(3.34) \quad = \underset{\lambda^{-2}\rho(\lambda)}{\text{Mean}} \frac{\iint_{-\infty}^{\infty} \tau\,(u, v\,;\, x, y\,;\, \lambda)\, du\, dv}{\iint_{-\infty}^{\infty} \tau_0(u, v\,;\, x, y\,;\, \lambda)\, du\, dv} .$$

Further,

$$\iint_{-\infty}^{\infty} \tau_0(u, v\,;\, x, y\,;\, \lambda)\, du\, dv =$$

$$= \left(\iint_{-\infty}^{\infty} |\, \mathcal{E}(u, v\,;\, x, y\,;\, \lambda)\,|\, du\, dv \right)^2$$

$$= \left(\iint_{-\infty}^{\infty} |\, \mathcal{E}_\lambda\,|\, du\, dv \right)^2 \Big/ \iint_{-\infty}^{\infty} |\, \mathcal{E}_\lambda|^2\, du\, dv ,$$

where $\mathcal{E}_\lambda$ is written for $\mathcal{E}(u, v\,;\, x, y\,;\, \lambda)$. Therefore

$$\frac{w(x, y\,;\, x, y)}{w_0(x, y\,;\, x, y)} =$$

$$= \underset{\lambda^{-2}\rho(\lambda)}{\text{Mean}} \left(\frac{\iint |\, \mathcal{E}_\lambda|^2\, du\, dv}{\left(\iint |\, \mathcal{E}_\lambda|\, du\, dv \right)^2} \iint_{-\infty}^{\infty} \tau(u, v\,;\, x, y\,;\, \lambda)\, du\, dv \right)$$

$$\geqq \underset{\lambda^{-2}\rho(\lambda)}{\text{Mean}} \left(\frac{|\, \mathcal{F}_{x, y\,;\, \lambda}\,|}{|\, \mathcal{C}\mathfrak{A}_{x, y\,;\, \lambda}\,|} \underset{\mathcal{F}_{x, y\,;\, \lambda}}{\text{Mean}}\, \tau(u, v\,;\, x, y\,;\, \lambda) \right)$$

(3.35)

$$= 4 \underset{\lambda^{-2}\rho(\lambda)}{\text{Mean}} \underset{\mathcal{F}_{x, y\,;\, \lambda}}{\text{Mean}}\, \tau(u, v\,;\, x, y\,;\, \lambda) ,$$

by Schwarz's inequality, with equality only if $|\, \mathcal{E}_\lambda|$ is constant over $\mathcal{C}\mathfrak{A}_{x, y\,;\, \lambda}$ for every (x, y) in F and every λ at which $\rho(\lambda) > 0$.

It follows from (3.35) that for any matching of image surface against object surface

$$\underset{\text{F}}{\text{Mean}} \frac{w(x, y\,;\, x, y)}{w_0(x, y\,;\, x, y)} \geqq 4 \underset{\lambda^{-2}\rho(\lambda)}{\text{Mean}} \underset{\text{F}}{\text{Mean}} \underset{\mathcal{F}_{x, y\,;\, \lambda}}{\text{Mean}}\, \tau(u, v\,;\, x, y\,;\, \lambda)$$

$$(3.36) \qquad = 4\, Q''_{\text{F}} ,$$

with equality only if $|\, \mathcal{E}_\lambda\,|$ is constant over $\mathcal{C}\mathfrak{A}_{x, y\,;\, \lambda}$. With an " optimal " matching, namely one which assigns the coordinates $x' = x$, $y' = y$ to an absolute maximum of $w(x', y'\,;\, x, y)$, the expression on the left of (3.36) becomes the extended Strehl definition S_{F} of the system in $\rho(\lambda)$-light [14] ; it follows that for any matching

$$(3.37) \qquad Q''_{\text{F}} \leqq \frac{1}{4}\, S_{\text{F}} ,$$

[14] It results from the definition of w_0 that $w_0(x, y\,;\, x, y) = \max_{(x', y')} w_0(x', y'\,;\, x, y)$, whatever the matching.

with equality only if the matching is optimal and $|\, \mathcal{E}_\lambda\,|$ is constant over $\mathcal{C}\mathfrak{A}_{x, y\,;\, \lambda}$. In systems of focal ratio $F/2$ or longer the last condition means that the absorption is substantially uniform over the aperture. It has a less simple meaning in systems of shorter focal ratio than $F/2$, which we do not consider here.

We can now state the result :

Whatever the choice of matching, the correlation quality in $\rho(\lambda)$-light of an unvignetted small-field system with rectangular or circular aperture, relative to a low-contrast random object set, is less than or equal to one quarter of the extended Strehl *definition. It can therefore never be made to exceed 1/4, even if aperture shading and central stopping are permitted. It can reach the value 1/4 only when* (1) *coordinate matching is optimised over F,* (2) *the absorption of the system is substantially uniform over the aperture,* (3) *the aberrations (including chromatism) are zero.*

Optimal matching for fidelity. — From the relation $2\, Q''_{\text{F}} = \Phi''_{\text{F}} + T''_{\text{F}}$, together with the invariance of T''_{F} under a small, smooth change of coordinate matching, it follows that the optimal matching for fidelity Φ''_{F} agrees with that for correlation quality Q''_{F}. When the object set is random under the contraint of a prescribed statistical mean effective spectral distribution $\rho(\lambda)$ this optimum matching assigns the coordinate numbers $x' = x$, $y' = y$ in the image surface to the " peak " (or point of maximum effective intensity) in the image of a point object situated at (x, y) in the object surface and emitting light of effective spectral distribution $\rho(\lambda)$. When the aberrations are small, this image has a single dominant peak, which moves " in step " with the object point (x, y) when the latter explores the field, and defines a coordinate mesh system $(x'\ y')$ with a smoothly varying metric. In systems with larger aberrations, a point source may give a multiple-peaked image whose isophote-geography alters as the point (x, y) moves from the centre to the edge of the field, and because there may be changes of leadership among the image peaks as the point (x, y) moves about the field, small abrupt changes in the optimum coordinate matching may occur on crossing certain boundary circles which divide the image surface into zones concentric with the optic axis. On the boundary circles themselves, the image of a point object has two or more highest peaks and consequently there are two or more choices of matching which give the fidelity and the correlation quality their respective greatest values. In small-field systems of uniform transmission and of focal ratio $F/2$ or longer, these greatest values are approximately given, in terms of the extended Strehl definition S_{F} and the relative structural content T''_{F}, by the equations

$$(3.38) \quad Q''_{\text{F max}} = \frac{1}{4}\, S_{\text{F}}, \quad \Phi''_{\text{F max}} = \frac{1}{2}\, S_{\text{F}} - T''_{\text{F}},$$

in which S_F and T_F'' are, to a good approximation, independent of the particular choice of matching.

We end this section by noting that the advantage of optimised correlation quality over extended Strehl definition as a concept is that it enables us to assess an optical system relative to an object set which is not fully random ; for example, relative to one which is random under the constraint of a prescribed statistical mean structural power spectrum. Only when this power spectrum is flat do the two concepts become equivalent, apart from a constant factor.

4. **Monochromatic image assessment on a receiving surface of known characteristics.** — An optical system is used with a receiving surface, and in some cases the properties of the receiving surface may safely be left out of account in assessing the quality of the system. Visual microscopes and astronomical telescopes used visually may be mentioned as example. In such cases, the assessments already discussed may supply a sufficiently realistic starting point for an analytical discussion of optical design problems.

Of at least equal practical importance are the cases in which the properties of the receiving surface must be explicitly taken into account in order to arrive at a realistic figure of merit for an optical system designed to form an image on that surface. Photographic instruments nearly all come under this head.

In the idealised special case where the light is substantially monochromatic and the " response " of the light-sensitive receiver is linear apart from random fluctuations, the response can be characterised over a small working field (say one of 15 degrees diameter or less) in terms of a spread-function $w_1(x', y')$ and a noise-function $n_2(x', y')$ in the following way :

The spread-function $w_1(x', y')$ may be defined as the statistical mean response to a luminous point source situated at $(x', y') = (0, 0)$ in the receiving surface S', normalised by the equation

$$(4.1) \qquad \iint_{-\infty}^{\infty} | w_1(x', y') | \mathrm{d}x'\mathrm{d}y' = 1 ;$$

we assume (what is true in most practical cases) that $w_1(x', y')$ is negligibly small unless the point (x', y') is very near to (0, 0). Uniformity of response from point to point of the receiving surface then corresponds analytically to the property that a point source situated at the point (x, y) in the image field F' gives, at an arbitrary point (x', y') in S', a response which can be taken as equal to $w_1(x' - x, y' - y)$ [15].

When the working field F is small the (x', y')-metric in S' is, for all (x, y) in F, approximately constant over the part of S' in which $w_1(x' - x, y' - y)$ is appreciably different from zero, and it then follows that the statistical mean response to the timage $I(x', y') = F^*[\tau F[\sigma]]$ of a single low-contrast object $\sigma(x, y)$ is represented, to a good approximation, by the function

$$(4.2) \qquad I(x', y') = w_1 * I,$$

where the convolution $w_1 * I$ is defined as usual by the equation

$$(4.3) \quad w_1 * I = \iint_{-\infty}^{\infty} w_1(x'' - x', y'' - y') I(x'', y'') \, \mathrm{d}x'' \, \mathrm{d}y''.$$

Taking Fourier transforms of both sides of (4.2), we obtain

$$(4.4) \qquad F[I_1] = \tau_1 F[I] = \tau\tau_1 F[\sigma],$$

where the " acceptance factor " $\tau_1(u, v)$ is defined by the equation

$$(4.5) \qquad \tau_1(u, v) = F[w_1]$$

and, by (4.1), $\tau_1(0, 0) = 1$.

By addition of the stochastic function $n_2(x', y')$ which represents the noise, we obtain from $I_1(x', y')$ the function

$$(4.6) \quad I_2(x', y') = I_1(x', y') + n_2(x, y') = F^*[\tau\tau_1 F[\sigma]] + n_2$$

which represents the response to a single presentation of the single object $\sigma(x, y)$.

When the response of the light-sensitive surface is not restricted to be linear, equation (4.6) still provides a statistically satisfactory approximate representation of $I_2(x', y')$ when the object $\sigma(x, y)$ is a member of a low-contrast random object set. The acceptance factor $\tau_1 = \tau_1(u, v)$ is now defined as $1/\tau_0(u, v)$ times the reduction in contrast in the image, by an aberration-free system of the same numerical aperture working with the given sensitive surface, of a low-contrast sinusoidal grating with spatial frequency (u, v) ; $\tau_0(u, v)$ here denotes the response factor of the aberration-free system. The spread function w_1 is defined by means of the equation

$$(4.7) \qquad W_1(x, y) = F^*[\tau_1],$$

which is equivalent to (4.5) by Fourier's inversion theorem. These definitions agree with the previous ones in the special case where the response is linear.

When aperture and field are given, a quality assessment of the optical system which takes account of the spread in the receiving surface can be obtained from each of T_F, Φ_F, and Q_F on replacing σ by $\sigma_* w_1 = \sigma_1$ and I by $I_* w_1 = I_1$ in the equations defining these quantities in the low-contrast case. To avoid complications, we consider first the case of a random low-

[15] Larger fields can be treated by the use of a spread-function $w_1(x'', y' ; x, y)$ but the formalism is less simple, since account then has to be taken of the effects of variation in (x', y')-metric over the field and of the obliquity of the principal ray of an off-axis image pencil. We therefore restrict the present discussion to fields in which these effects can be disregarded without serious inaccuracy.

contrast object set and a substantially noise-free receiving surface. Making use of the relations

$$F[\sigma_* w_1] = F[\sigma]\, F[w_1] = \varepsilon(u, v)\, \tau_1(u, v)$$
$$F[I_* w_1] - F[I]\, F[w_1] \sim \tau(u, v; x, y)\, \tau_1(u, v)\, \varepsilon(u, v),$$

valid in an isoplanatism patch surrounding (x, y), we obtain after calculation the structural assessments

$$(4.8) \quad T_{\mathrm{F}}''', \Phi_{\mathrm{F}}''', Q_{\mathrm{F}}''' = \underset{\mathrm{F},\,\mathcal{F}}{\mathrm{Mean}} |\tau_1|^2\, \Psi(\tau) \Big/ \underset{\mathcal{F}}{\mathrm{Mean}} |\tau_1|^2$$

as generalisations of (3.16), (3.18), to which they reduce formally when the spread is negligibly small, i.e. when $\tau_1(u, v) = 1$ for all (u, v) in $\mathcal{F}$. Thus T_{F}''', Φ_{F}''' and Q_{F}''' are the $|\tau_1|^2$-weighted $(u, v; x, y)$-means of $|\tau|^2$, $1 - |1 - \tau|^2$ and τ [or $\mathcal{R}(\tau)$] respectively, where (x, y) runs over F and (u, v) over the region $\mathcal{F}$ in the (u, v)-plane.

An optical design which maximises one of the expressions (4.8) gives the highest mathematical expectation of image quality in the corresponding sense when used, with a receiving surface of spread w_1 and negligible noise, to observe a previously unknown low-contrast object set ([16]).

When the low-contrast object set in F has maximum randomness under the constraint of a prescribed statistical mean power spectrum $\varphi(u, v)$, we obtain in place of T_{F}''', Φ_{F}''', Q_{F}''' the structural assessments

$$(4.9) \quad T_{\mathrm{F},\varphi}''', \Phi_{\mathrm{F},\varphi}''', Q_{\mathrm{F},\varphi}''' =$$

$$= \mathrm{Mean}\left(\int\!\!\int_{-\infty}^{\infty} \Psi(\tau)\, |\tau_1|^2 \varphi\, du\, dv \Big/ \int\!\!\int_{-\infty}^{\infty} |\tau_1|^2 \varphi\, du\, dv \right)$$

of image-quality. These assessments are the $(|\tau_1|^2\varphi)$-weighted $(u, v; x, y)$-means of $|\tau|^2$, $1 - |1 - \tau|^2$ and τ [or $\mathcal{R}(\tau)$] respectively, where (x, y) runs over F and (u, v) over the whole (u, v)-plane. The corresponding optimal image qualities are obtained on maximising $T_{\mathrm{F},\varphi}'''$, $\Phi_{\mathrm{F},\varphi}'''$ and $Q_{\mathrm{F},\varphi}'''$ by variation of the available design parameters. If aperture as well as field is fixed in advance, this is equivalent to maximising the $(|\tau_1|^2 \varphi)$-weighted means over $\mathrm{F} \times \mathcal{R}$ of $|\tau|^2$, $1 - |1 - \tau|^2$ and $\mathcal{R}(\tau)$ respectively.

The $|\tau_1|^2$-weighting has a powerful effect on the optimum design of a high-aperture photographic objective. In an $F/2$ lens used with a photographic emulsion of resolution limit 0.025 mm, the values of (u, v) for which $|\tau_1|^2$ differs significantly from zero are all contained in a frequency-circle $\mathcal{F}_1$, concentric with $\mathcal{F}$ and of $\frac{1}{10}$ the radius. Only in this circle of frequencies can a variation in τ affect the photographic image appreciably. In balancing the aberrations to the best advantage, we can therefore afford to ignore the behaviour of τ at frequencies outside the circle $\mathcal{F}_1$. Photographic tests with low-contrast sinusoidal gratings as test objects provide a convenient practical means of measuring $\tau\tau_1$.

The structural assessments (4.9) can be extended without essential change to cover the case where image-noise is no longer negligible. A brief indication of the procedure will perhaps be sufficient. In effect, the image-intensity function I_1 is replaced by $I_2 = I_1 + n_2$ [see (4.6)] in the integrals

$$\int\!\!\int_{\mathrm{F}} \langle\, I_1^2 \,\rangle\, dx\, dy, \qquad \int\!\!\int_{\mathrm{F}} \langle\, (\sigma_1 - I_2)^2 \,\rangle\, dx\, dy,$$
$$\int\!\!\int_{\mathrm{F}} \langle\, \sigma_1\, I_1 \,\rangle\, dx\, dy$$

on which the equations for T, $1 - \Phi$ and Q are ultimately based, and the statistical means are now taken over object-set and image noise simultaneously. In the first assessment, that part of the mean structural content which arises from the noise n_2 is regarded as spurious and subtracted off. The background intensities $B_\sigma\, \varkappa_{\mathrm{F}}(x, y)$ are subtracted off, as before [compare (3.13), (3.14)] in each case. It then turns out that the image quality assessment in each case depends on the noise only through the total noise power, while the corresponding optimum design, though dependent on the prescribed field-size, aperture and focal length, on the image-spread in the receiving surface, and on the statistics of the presumed low-contrast object set, is independent of the noise.

Now the resolution limit in the receiving surface is essentially governed by the relation between $|\tau_1(u, v)|^2$ and the power density of the noise spectrum at the frequency (u, v). For example, the statement that the resolution limit of a photographic emulsion is 0.025 mm means that at spatial frequencies higher than 40 mm^{-1} the value of $|\tau_1|^2$ is so small that image detail is completely smothered in noise. An increase in noise-level, without appreciable change in τ_1, results in some frequencies below 40 mm^{-1} being completely smothered. Then the domain of frequencies which need to be considered in optimising the optical design from the point of view of distinguishing image detail is smaller than before, and consequently the optimum design from this point of view will be changed when the noise level is changed.

The fact that, on each of the extended assessments described above, the optimised design is not affected by a change in the noise appears natural enough when we observe that what these assessments measure is the average similarity (in some sense) of image to object, not the capacity of the system to produce mu-

([16]) Lack of previous knowledge is represented analytically by assigning to the object set maximum randomness, in the sense of maximum entropy. In the present case this corresponds (in the sense of § 3 above) to a set of uniform statistical mean spectral power distribution in $\mathcal{F}$.

tually distinguishable images of the objects in a low-contrast set. An assessment which measures this capacity, and does not explicity concern itself with the similarity of image to object, is the assessment by mean information content, discussed in a recent joint paper with P. B. Fellgett (1955). In this assessment, the noise plays a fundamental part. For example, if it is desired to resolve as efficiently as possible the fine detail in the noisy images of a low-contrast object set, random except for a prescribed statistical mean effective structural power spectrum $\varphi(u, v)$, on a receiving surface with gaussian noise of known statistical mean power spectrum $N(u, v)$, the design may be informationally optimised by maximising the mean information content

$$(4.10) \quad |\mathrm{F}\|\mathfrak{F}| \underset{\mathrm{F},\mathfrak{F}}{\mathrm{Mean}} \log \sqrt{\left(1 + \frac{\varphi(u, v)}{N(u, v)} |\tau\tau_1|^2\right)}$$

of the images of this object set.

It results from the special properties of centred optical systems with spherical or figured surfaces, and with the usual simple forms of aperture, that a system of high informational quality also has a high correlation quality, a high structural ratio and a high fidelity. In particular, the structural assessment (3.16) of a system imaging a random low-contrast object set of prescribed statistical mean total structure power P and photon noise-power N agrees, except for a normalising factor, with its informational assessment (4.10) in the limiting case where $P \ll N$, i. e. where structure is almost completely smothered in noise. In the general case, analogous but less simple relations result from the fact that the image of a point object cannot be " sharp " [in the sense of large equivalent band-radius ([17])] unless it is also small. Of course, this result need not hold for systems with more complicated types of aperture or of wave-aberration, such as telescopes fitted with wire objective gratings, or with grating spectrographs. In these instruments the informational assessment of the system is an appropriate one, but the fidelity assessment obviously gives no indication of the optical quality. As remarked in section 1, the assessment must be chosen to match the purpose for which the system is intended.

([17]) The importance of this quantity as a means of measuring image sharpness was pointed out in an earlier paper (Linfoot, 1957). The essential idea goes back to Schade (1952) and Fellgett (1953).

References

J. W. Coltman, *J. O. S. A.*, **44**, 468, 1954.
P. B. Fellgett, *J. O. S. A.*, **43**, 271, 1953.
P. B. Fellgett and E. H. Linfoot, *Phil. Trans. Roy. Soc.* (A), **247**, 369, 1955.
H. H. Hopkins, *Proc. Roy. Soc.* (A), **231**, 91, 1955.
P. Lindberg, *Optica Acta*, **1**, 80, 1954.
E. H. Linfoot, *J. O. S. A.*, **45**, 808, 1955, *J. O. S. A.*, **46**, 740 1956, *Optica Acta*, **4**, 12, 1957.
O. H. Schade, *J. Soc. Motion Picture and Television Engrs.*, **58**, 182, 1952.
K. Strehl, *Z. Instrumentenkunde*, **22**, 213, 1902.
F. E. Washer, Contr. to " Optical Image Evaluation " (Washington, 1954), 209.

Manuscrit reçu le 1er juillet 1957.

Influence de l'éclairage partiellement cohérent sur la formation des images de quelques objets étendus opaques.

D. Canals-Frau et Mme M. Rousseau,
Institut d'Optique, Paris

Sommaire. — *On établit les formules donnant l'intensité et le contraste des images optiques de quelques objets particuliers éclairés en lumière partiellement cohérente. On étudie les images de la ligne opaque, du bord de plage et du disque noir en fonction du degré de cohérence de l'éclairage. On suppose toujours l'instrument dépourvu d'aberrations et la source parfaitement monochromatique.*

Zusammenfassung. — *Die Formeln für die Intensität und den Kontrast der optischen Abbildungen einiger besonderer Objekte bei Beleuchtung mit teilweise kohärentem Licht werden aufgestellt. Die Abbildungen der undurchsichtigen geraden Linie, der Kante und der dunklen Scheibe werden in ihrer Abhängigkeit von dem Kohärenzgrad der Beleuchtung untersucht. Dabei wird stets das abbildende System als aberrationsfrei und das Licht als rein monochromatisch vorausgesetzt.*

Summary. — *Formulae are obtained for the intensity and contrast of the optical images of certain objects formed by partially coherent light. The images of an opaque line, an edge and an opaque disc are considered as a function of the degree of coherence of the illumination. Imagery free from aberration and in monochromatic light is assumed in all cases.*

1. **Introduction.** — Dans l'« Optique classique », on fait une distinction rigide entre éclairage cohérent et éclairage incohérent : dans le premier cas, on obtient l'intensité de l'image en additionnant les amplitudes complexes puis en prenant le carré du module ; dans le second cas, on additionne les carrés des modules des amplitudes complexes.

Récemment, plusieurs auteurs, notamment Zernike [1], Hopkins [2,3], Duffieux [4], Blanc-Lapierre et Dumontet [5], Wolf [6] ont introduit la théorie de la cohérence partielle.

Dans le cas particulier de l'Optique, le degré de cohérence est déterminé uniquement par des facteurs géométriques : l'angle sous lequel l'instrument « voit » la source. En microscopie, par exemple, il est bien connu qu'il faut fermer le condenseur pour obtenir de bonnes figures d'interférences. On est amené à se demander comment varie l'aspect de l'image en fonction de l'ouverture du condenseur. Maréchal [7] a montré que, dans le cas particulier d'un petit objet noir sur fond blanc (point ou ligne), le contraste doublait quand on passait de l'éclairage incohérent à l'éclairage cohérent.

Dans le présent travail, nous nous proposons de déterminer, en fonction du degré de cohérence de l'éclairage, les images de quelques objets étendus opaques. Nous donnons le contraste de l'image d'un ruban noir et d'un disque noir et l'éclairement de l'image d'un demi-plan et d'une ligne noire.

Nous supposons l'instrument dépourvu d'aberration et la lumière parfaitement monochromatique.

Les calculs sont faits en admettant la linéarité des filtres (au sens de Blanc-Lapierre et Dumontet [5], c'est-à-dire en supposant que la tache de diffraction reste invariable (en module et en argument) quand on translate le point-objet dans le champ.

Nous utilisons, comme formules de départ, les expressions données par Hopkins [2, 3].

2. **Notations et hypothèses.** — La figure 1 représente le schéma très simplifié d'un système optique, centré (par exemple un microscope) d'axe zz'.

Le condenseur, dont le plan pupillaire est indiqué, sur la figure, par P_c, possède dans l'espace-objet l'ouverture numérique $n_c \sin \alpha_c$. De même l'objectif

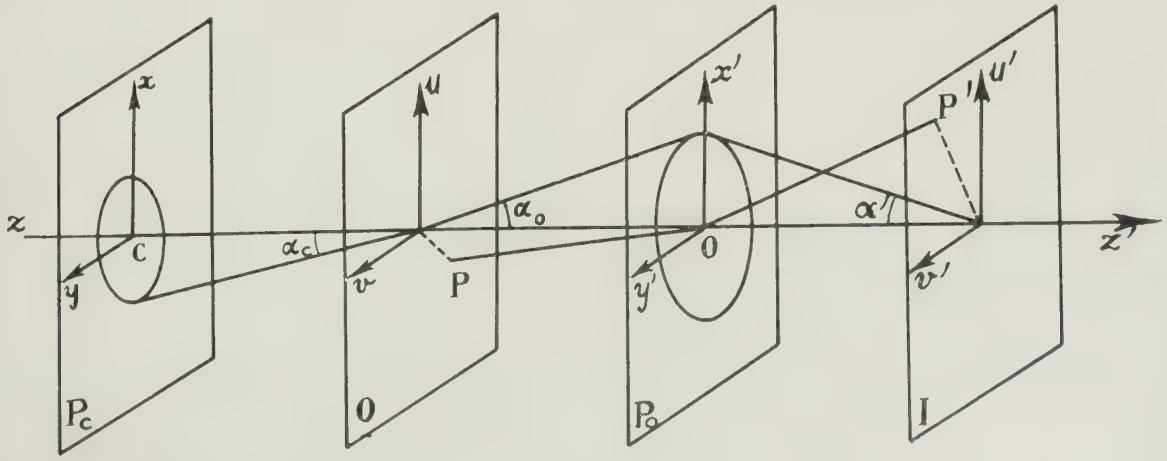

Fig. 1.

a pour pupille de sortie (P_0) et pour ouverture numérique $n_0 \sin \alpha_0$.

Soit M un point du plan objet. Soient u, v, les coordonnées réduites de M reliées à ses coordonnées géométriques ξ, η par le facteur

$$(1) \qquad \frac{2\pi}{\lambda} n_0 \sin \alpha_0 = \frac{2\pi}{\lambda} n_c \sin \alpha_c$$

telles que [3]

$$u = \frac{2\pi}{\lambda} n_0 \sin \alpha_0 \, \xi ; \qquad v = \frac{2\pi}{\lambda} n_0 \sin \alpha_0 \, \eta.$$

De même, dans le plan image,

$$u' = \frac{2\pi}{\lambda} n' \sin \alpha' \, \xi' ; \qquad v' = \frac{2\pi}{\lambda} n' \sin \alpha' \, \eta'.$$

Avec ce choix de coordonnées, le point P′ de coordonnées ($u' = u$, $v' = v$) est l'image géométrique du point P(u, v).

Si on prend pour unité de longueur l'inverse de la fréquence spatiale limite.

$$(2) \qquad \rho = \frac{\lambda}{2 n_0 \sin \alpha_0} ,$$

on a

$$u = \pi U, \qquad v = \pi V ;$$

U et V sont alors les coordonnées géométriques de l'objet mesurées avec ρ comme unité de longueur,

$$(3) \qquad U = \frac{\xi}{\rho} , \qquad V = \frac{\eta}{\rho} .$$

Les coordonnées polaires correspondantes à (u, v) et (u', v') seront désignées par (ρ, θ) et (ρ', θ'). Un

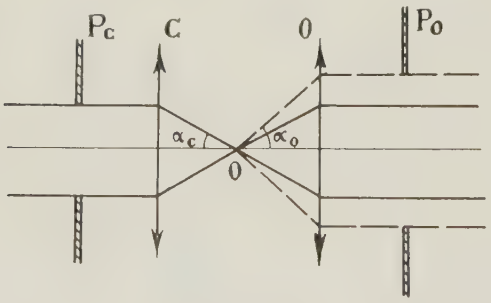

Fig. 2 a.

point M de la pupille du condenseur sera défini en coordonnées cartésiennes par x, y et en coordonnées polaires par r et φ (en prenant pour unité de longueur la moitié de la largeur de la pupille du condenseur) : Un point M′ de la pupille de l'objectif sera caractérisé par x', y' ou r', φ' (en prenant pour unité la demi-largeur de la pupille de l'objectif).

L'éclairage d'un microscope peut être caractérisé par le rapport :

$$(4) \qquad s = \frac{n_c \sin \alpha_c}{n_0 \sin \alpha_0} \qquad \text{(fig. 2 a).}$$

La figure 2b représente le plan de la pupille de sortie (P_0) de l'objectif sur laquelle vient se former l'image (P_c') de la pupille (P_c) du condenseur.

Pour simplifier l'étude de certains objets : demi-plan, fils..., nous avons été amenés à considérer un instrument équipé de pupilles en fentes.

En supposant les fentes très longues par rapport à

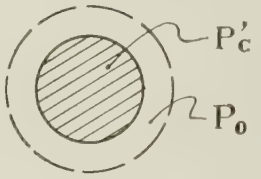

Fig. 2 b.

leur largeur, on peut ramener l'étude à un problème unidimensionnel. La figure 3 est, comme précédemment, un rabattement de la pupille de l'objectif mettant en évidence le rapport s.

Pour les objets dont les dimensions n'excèdent pas

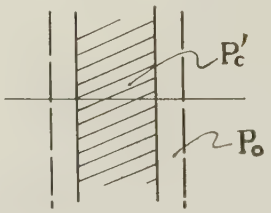

Fig 3.

celles de la tache de diffraction, nous avons pris, comme définition du contraste,

$$(5) \qquad C = \frac{I(\infty) - I(0)}{I(\infty)} ,$$

où $I(0)$ est l'intensité au centre de l'image
et $I(\infty)$ est l'intensité très loin du bord géométrique de l'image.

3. **Pupilles en fentes.** — L'instrument est équivalent à un système optique à symétrie cylindrique dans lequel on suppose que le condenseur, l'objectif et l'objet s'étendent indéfiniment dans la direction des axes x et u (tous les plans perpendiculaires aux axes x et u sont équivalents). Dans ces conditions, les formules (si l'objet présente la même symétrie) ne dépendent plus que d'une seule variable.

L'éclairement de l'image se réduit à

$$(6) \qquad \mathfrak{I}(v') = A \int_{-\infty}^{+\infty} \gamma(y) \left| \Phi(v', s\, y) \right|^2 \mathrm{d}y$$

où $\gamma(y)$ est la distribution d'éclairement sur la pupille du condenseur que nous pouvons prendre, sans perte de généralité,

$$(7) \qquad \gamma(y) = \begin{cases} \gamma \text{ pour } |y| < 1 \\ 0 \text{ pour } |y| > 1, \end{cases}$$

et

$$(8)\quad \Phi(\varphi', sy) = \int_{-\infty}^{+\infty} O(\varphi)\,.\,F(\varphi' - \varphi)\,\mathrm{e}^{+\mathrm{i}sy\varphi}\,\mathrm{d}\varphi$$

$$(9)\quad F(\varphi' - \varphi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(y)\,\mathrm{e}^{-\mathrm{i}(\varphi' - \varphi)y}\,\mathrm{d}y$$

est le filtre linéaire unidimensionnel qui caractérise l'objectif : c'est l'amplitude complexe produite en u' par un point source (d'amplitude unité et de phase nulle) placé en u. Alors $f(y)$ est la transmission complexe de la pupille ; dans le cas présent, en l'absence d'aberration et en supposant l'absorption nulle, on prendra

$$(10)\qquad f(y) = \begin{cases} 1 \text{ pour } |y| < 1 \\ 0 \text{ pour } |y| > 1, \end{cases}$$

$O(\varphi)$ est la transmission complexe de l'objet.

A. ***Image d'un ruban opaque.*** — Soit

$$2\,\varepsilon_0 = \frac{4\pi}{\lambda}\, n_0 \sin \alpha_0\, \eta_0$$

la largeur de ce ruban (fig. 4)....

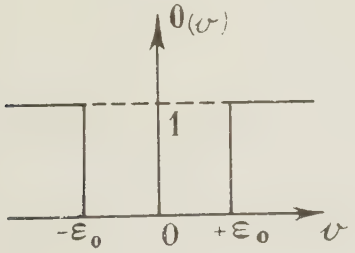

Fig. 4.

Dans ce cas,

$$(11)\qquad O(\varphi) = \begin{cases} 1 \text{ pour } |\varphi| > \varepsilon_0 \\ 0 \text{ pour } |\varphi| < \varepsilon_0. \end{cases}$$

La formule (8) se réduit à

$$(12)\quad \Phi(\varphi', sy) = \int_{-\infty}^{+\infty} F(\varphi' - \varphi)\,\mathrm{e}^{+\mathrm{i}sy\varphi}\,\mathrm{d}\varphi - \int_{-\varepsilon_0}^{+\varepsilon_0} F(\varphi' - \varphi)\,\mathrm{e}^{+\mathrm{i}sy\varphi}\,\mathrm{d}\varphi.$$

Le 1er terme du 2e membre devient, en utilisant la transformée inverse de (9),

$$\sqrt{2\pi}\,\mathrm{e}^{\mathrm{i}sy\varphi'}\, f(-sy).$$

Dans la 2e intégrale, il faut remplacer $F(\varphi' - \varphi)$ par sa valeur

$$(13)\qquad F(\varphi' - \varphi) = \sqrt{\frac{2}{\pi}}\,\frac{\sin(\varphi' - \varphi)}{(\varphi' - \varphi)}\,.$$

On obtient

$$(14)\quad \int_{-\varepsilon_0}^{+\varepsilon_0} F(\varphi' - \varphi)\,\mathrm{e}^{+\mathrm{i}sy\varphi}\,\mathrm{d}\varphi = \sqrt{\frac{2}{\pi}}\,\mathrm{e}^{\mathrm{i}sy\varphi'} \int_{\varphi' - \varepsilon_0}^{\varphi' + \varepsilon_0} \frac{\sin w}{w}\,\mathrm{e}^{-\mathrm{i}syw}\,\mathrm{d}w$$

$$- \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}sy\varphi'} \times \times \left[\int_0^{\varphi' + \varepsilon_0} \frac{\sin w(1+ys) + \sin w(1-ys)}{w}\,\mathrm{d}w - \int_0^{\varphi' - \varepsilon_0} \frac{\sin w(1+ys) + \sin w(1-ys)}{w}\,\mathrm{d}w + \mathrm{i} \left\{ \int_{\varphi' + \varepsilon_0}^{+\infty} \frac{\cos w(1-ys) - \cos w(1+ys)}{w}\,\mathrm{d}w - \int_{-\infty}^{\varphi' - \varepsilon_0} \frac{\cos w(1-ys) - \cos w(1+ys)}{w}\,\mathrm{d}w \right\} \right]$$

$$= \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}sy\varphi'} \left[H(\varphi', sy, \varepsilon_0) - \mathrm{i}\,K(\varphi', sy, \varepsilon_0) \right]$$

en posant

$$(15)\quad H = \Big\{ \mathrm{Si} \Big\} (\varphi' + \varepsilon_0)(1 + ys) \Big\{ + \mathrm{Si} \Big\} (\varphi' + \varepsilon_0)(1 - ys) \Big\{ - \mathrm{Si} \Big\} (\varphi' - \varepsilon_0)(1 + ys) \Big\{ - \Big\} \mathrm{Si}(\varphi' - \varepsilon_0)(1 - ys) \Big\{ \Big]$$

$$(16)\quad K = \Big[- \mathrm{Ci} \Big\{ |\varphi' - \varepsilon_0|\,(1 - ys) \Big\} + \Big\} \mathrm{Ci}\,|\varphi' - \varepsilon_0|\,(1 + ys) \Big\} + \Big\{ \mathrm{Ci}\,|\varphi' + \varepsilon_0|\,(1 - ys) \Big\} - \Big\} \mathrm{Ci}\,|\varphi' + \varepsilon_0|\,(1 + ys) \Big\} \Big]$$

l'équation (12) s'écrit

$$\Phi(\varphi', sy) = \sqrt{2\pi}\,\mathrm{e}^{\mathrm{i}sy\varphi'}\, f(-sy) - \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{\mathrm{i}sy\varphi'}\,[H - \mathrm{i}\,K],$$

$$(17)\quad |\Phi(\varphi', sy)|^2 = 2\pi\,[f(-sy)]^2 - 2\,H\,f(-sy) + \frac{1}{2\pi}\,[H^2 + K^2]$$

d'où

$$(18)\quad \mathcal{J}(\varphi') = 2\,A \int_{-\infty}^{+\infty} \Big\{ \pi\,\gamma(y)\,|f(-sy)|^2 - \gamma(y) \times H \times f(-sy) + \frac{1}{4\pi}\,\gamma(y)\,(H^2 + K^2) \Big\}\,\mathrm{d}y\,.$$

En raison de la définition de $\gamma(y)$, équation (7), et de $f(-sy)$, équation (10), il faut distinguer les cas $s > 1$ et $s < 1$.

Pour $s < 1$, sy est toujours < 1, c'est la fonction $\gamma(y)$ qui impose les limites d'intégration.

Pour $s > 1$, $f(-sy) = f(sy) = 1$ seulement pour les valeurs de y inférieures à $\frac{1}{s}$; c'est la fonction $f(sy)$ qui limite l'intégrale.

Les 2 derniers termes de l'intégrale sont les mêmes dans les 2 cas.

a) $s < 1$. — On normalisera la fonction (18) en la divisant par l'éclairement qui existe dans le plan image en l'absence de l'objet, c'est-à-dire

$$(19) \qquad \mathfrak{I}(\varphi', \varepsilon_0 = 0, s) = 4\pi A\gamma .$$

L'éclairement normalisé est

$$(20) \qquad I(\varphi', \varepsilon_0, s) = \frac{\mathfrak{I}(\varphi', \varepsilon_0, s).}{\mathfrak{I}(\varphi', \varepsilon_0 = 0, s)} .$$

L'intensité au point φ' de l'image est

$$(21) \qquad I(\varphi') = \frac{1}{2\pi\gamma} \int_{-1}^{+1} \left\{ \pi\gamma - \gamma H + \right.$$
$$\left. + \frac{\gamma}{4\pi} (H^2 + K^2) \right\} dy$$
$$= X_1 + Y + Z$$

avec

$$(22) \quad X_1(\varphi', \varepsilon_0, s) = \frac{1}{2\pi\gamma} \int_{-1}^{+1} (\pi\gamma - \gamma H)\, dy,$$
$$Y(\varphi', \varepsilon_0, s) = \frac{1}{2\pi\gamma} \int_{-1}^{+1} \frac{\gamma}{4\pi} H^2\, dy,$$
$$Z(\varphi', \varepsilon_0, s) = \frac{1}{2\pi\gamma} \int_{-1}^{+1} \frac{\gamma}{4\pi} K^2\, dy.$$

On calcule séparément les expressions X_1, Y, Z.

En posant

$$(23) \qquad a = \varphi' + \varepsilon_0, \quad b = \varphi' - \varepsilon_0.$$

l'expression de l'éclairement devient

$$(24) \qquad I(\varphi') =$$
$$= 1 - \frac{1}{\pi s} \left[(1+s) \left\{ \mathrm{Si}\, a(1+s) - \mathrm{Si}\, b(1+s) \right\} - \right.$$
$$- (1-s) \left\{ \mathrm{Si}\, a(1-s) - \mathrm{Si}\, b(1+s) \right\} -$$
$$\left. - 2 \frac{\sin a \,.\, \sin sa}{a} + 2 \frac{\sin b \,.\, \sin sb}{b} \right] +$$
$$+ \frac{1}{4\pi^2 s} \left\{ \mathfrak{F}(s, a) - \mathfrak{F}(-s, a) + \mathfrak{F}(s, b) - \mathfrak{F}(-s, b) + \right.$$
$$\left. + \mathfrak{F}_1(s, b) - \mathfrak{F}_1(-s, b) + \mathfrak{F}_1(s, a) - \mathfrak{F}_1(-s, a) \right\} +$$
$$\frac{1}{2\pi^2} \int_0^{+1} \Big[G(a, s\,;\, a, -s\,;\, y) - G(a, s\,;\, b, s\,;\, y) -$$
$$- G(a, s\,;\, b, -s\,;\, y) - G(a, -s\,;\, b, s\,;\, y) +$$
$$- G(a, -s\,;\, b, -s\,;\, y) + G(b, s\,;\, b, -s\,;\, y) -$$
$$- G_1(b, -s\,;\, b, s\,;\, y) - G_1(b, -s\,;\, a, -s, y) +$$
$$+ G_1(b, -s\,;\, a, s\,;\, y) + G_1(b, s\,;\, a, -s\,;\, y) -$$
$$- G_1(b, s\,;\, a, s\,;\, y) - G_1(a, -s\,;\, a, s\,;\, y) \Big] dy$$

où

$$(25) \qquad \mathfrak{F}(s, a) = (1+s)\, \mathrm{Si}^2 [\, a(1+s)\,] +$$
$$+ 2 \frac{\cos[\, a(1+s)\,] \times \mathrm{Si}[\, a(1+s)\,]}{a} - \frac{\mathrm{Si}\, 2a(1+s)}{a} ,$$

$\mathfrak{F}_1$ a la même forme que la formule $\mathfrak{F}$; dans les 2 premiers termes, on a remplacé Si par — Ci et cosinus par sinus.

$$(26) \quad G(a, s\,;\, a, -s\,;\, y) = \mathrm{Si}[\, a(1+sy)\,] \times$$
$$\times \mathrm{Si}[\, a(1-sy)\,] .$$

On obtient G_1 en remplaçant Si par Ci dans G.

Pour de faibles ouvertures et des rubans minces, on peut conserver seulement les 2 premiers termes de (24).

$$(27) \quad I(\varphi', \varepsilon_0) = 1 - \frac{2}{\pi} \Big[\mathrm{Si}(\varphi' + \varepsilon_0) - \mathrm{Si}(\varphi' - \varepsilon_0) \Big] .$$

Pour un point φ' donné, l'éclairement I est fonction de ε_0. La quantité ε_0 est égale au produit de la demi-largeur de la ligne par l'ouverture de l'objectif. Pour faire varier ε_0, on peut agir indépendamment sur l'un ou l'autre de ces facteurs.

Pour $\varepsilon_0 < 1$, on retrouve l'expression de la cohérence.

$$I(\varphi', \varepsilon_0) = 1 - \frac{4\,\varepsilon_0}{\pi} \frac{\sin \varphi'}{\varphi'} .$$

Eclairement de l'image d'un ruban opaque en lumière cohérente. — Cherchons la valeur que prend l'éclairement quand s tend vers zéro.

$$(28) \qquad \lim_{s \to 0} I(\varphi', \varepsilon_0, s) = I_c(\varphi', \varepsilon_0)$$
$$= 1 - \frac{2}{\pi} \Big[\mathrm{Si}\, a - \mathrm{Si}\, b \Big] + \frac{1}{\pi^2} \Big[\mathrm{Si}\, a - \mathrm{Si}\, b \Big]^2 =$$
$$= \left\{ 1 - \frac{1}{\pi} \Big[\mathrm{Si}(\varphi' + \varepsilon_0) - \mathrm{Si}(\varphi' - \varepsilon_0) \Big] \right\}^2 .$$

b) $s \geqslant 1$. — Comme précédemment, pour normaliser, nous diviserons l'éclairement (18) par

$$(29) \quad I(\varphi', \varepsilon_0 = 0, s) = 2A \int_{-1/s}^{+1/s} \pi\gamma\, dy = \frac{4\pi A\, \gamma}{s} .$$

De telle sorte que

$$I(\varphi') = \frac{s}{2\pi\gamma} \int_{-1/s}^{+1/s} (\pi\gamma - \gamma H)\, dy +$$
$$+ \frac{s}{8\pi^2} \int_{-1}^{+1} (H^2 + K^2)\, dy .$$

Il faut calculer à nouveau les deux premiers termes de cette intégrale. Les deux derniers termes, non affectés par le facteur f, gardent la même valeur que dans le cas $s < 1$.

L'éclairement s'écrit

$$I(\varphi') = X_2 + s(Y + Z)$$

avec

$$X_2(\varphi', \varepsilon_0, s) = \frac{s}{2\pi\gamma} \int_{-1/s}^{+1/s} (\pi\gamma - \gamma H)\, dy .$$

La formule générale de l'éclairement dans le cas d'un rapport d'ouverture dépassant l'unité est

$$(30) \quad I(\varphi') =$$
$$= 1 - \frac{2}{\pi} \left\{ \mathrm{Si}\, 2a - \mathrm{Si}\, 2b - \frac{\sin^2 a}{a} + \frac{\sin^2 b}{b} \right.$$

$$+\frac{1}{4\pi^2}\Big\{\mathcal{F}(s,a)-\mathcal{F}(-s,a)+F(s,b)-\mathcal{F}(-s,b)$$
$$+\mathcal{F}_1(s,b)-\mathcal{F}_1(-s,b)+\mathcal{F}_1(s,a)+\mathcal{F}_1(-s,a)\Big\}+$$
$$+\frac{1}{2\pi^2}\int_0^s\Big[G(a,1\,;\,a,-1,y)-G(a,1\,;\,b,1,y)-$$
$$-G(a,1\,;\,b,-1\,;\,y)-G(a,-1\,;\,b,1\,;\,y)-$$
$$-G(a,-1\,;\,b,-1\,;y)+G(b,1\,;\,b,-1\,;\,y)-$$
$$-G_1(b,-1\,;\,b,1\,;\,y)-G_1(b,-1\,;\,a,-1\,;\,y)+$$
$$+G_1(b,-1\,;\,a,+1\,;\,y)+G_1(a,1\,;\,a,-1\,;\,y)-$$
$$-G_1(b,1\,;\,a,1\,;\,y)-G_1(a,-1\,;\,a,1\,;\,y)\Big]\,dy\,;$$

$\mathcal{F}$, $\mathcal{F}_1$, G et G_1 ont la même signification que plus haut.

Eclairement dans le cas d'un éclairage incohérent. — Il faut revenir à la formule initiale (6).

$$\mathfrak{I}(\varphi')=A\int_{-\infty}^{+\infty}\gamma(y)\,\Phi(\varphi',sy)\times\Phi^*(\varphi',sy)\,dy$$
$$=A\int_{-\infty}^{+\infty}\gamma(y)\int_{-\infty}^{+\infty}O(\varphi)\,F(\varphi'-\varphi)\,e^{isy\varphi}\,d\varphi$$
$$\int_{-\infty}^{+\infty}O^*(w)\,F^*(\varphi'-w)\,e^{-isyw}\,dw$$
$$=A\,\gamma\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}O(\varphi)\,O^*(w)\,F(\varphi'-\varphi)\,F^*(\varphi'-w)$$
$$\frac{1}{s}\int_{-s}^{+s}e^{i\beta(\varphi-w)}\,d\beta\,d\varphi\,dw\,.$$

En normalisant cette équation avec la valeur (29) on a, pour le cas incohérent (s tendant vers l'infini),

$$I(\varphi',\varepsilon_0)=$$
$$\lim_{s\to\infty}=\lim_{s\to\infty}\frac{1}{4\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}O(\varphi)\,O^*(w)\,F(\varphi'-\varphi)\,F^*(\varphi'-w)$$
$$\int_{-s}^{+s}e^{i\beta(\varphi-w)}\,d\beta\,d\varphi\,dw=I_i(\varphi',\varepsilon_0)\,.$$

En utilisant la distribution de Dirac,

$$\lim_{s\to\infty}\int_{-s}^{+s}e^{i\beta(\varphi-w)}\,d\beta=2\pi\,\delta(\varphi-w)\,,$$

on obtient

$$(28')\quad I_i(\varphi',\varepsilon_0)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}|O(\varphi)|^2\times|F(\varphi'-\varphi)|^2\,d\varphi$$
$$=1-\frac{1}{\pi}\Big\{\mathrm{Si}\,[2(\varphi'+\varepsilon_0)]-\mathrm{Si}\,[2(\varphi'-\varepsilon_0)]-$$
$$-\frac{\sin^2(\varphi'+\varepsilon_0)}{(\varphi'+\varepsilon_0)}+\frac{\sin^2(\varphi'-\varepsilon_0)}{(\varphi'-\varepsilon_0)}\Big\}.$$

Cette formule est la limite naturelle que prend l'équation générale (6) quand s tend vers l'infini. De même, l'équation (28) est obtenue comme limite de (6) quand s tend vers zéro, à condition toutefois d'effectuer les normalisations dès le départ. En effet, la normalisation pour $s>1$ est différente de la normalisation pour $s<1$. C'est pourquoi les images en lumière cohérente et en lumière incohérente proviennent finalement de deux équations différentes. Toutefois, les deux formules générales pour $s<1$ et $s>1$ coïncident pour la valeur $s=1$, mais leurs dérivées ne coïncident pas.

B. ***Contraste de l'image d'un ruban opaque.*** — Dans le cas d'un ruban de largeur inférieure à l'inverse de la fréquence spatiale limite le contraste peut être exprimé par

$$C=\frac{I(\infty)-I(0)}{I(\infty)}$$

(dans le cas où la bande est beaucoup plus large, ce contraste est toujours égal à l'unité).

a) $s<1$. — L'éclairement au centre de l'image est

$$I(0)=\frac{1}{2\pi\gamma}\Big[X_1(0,\varepsilon_0,s)+Y(0,\varepsilon_0,s)+Z(0,\varepsilon_0,s)\Big]$$

on a, maintenant, $a=-b=\varepsilon_0$.

$Z(0,\varepsilon_0,s)$ se réduit à zéro puisque K s'annule pour $\varphi'=0$ (16),

$$I(0)=1-\frac{2}{\pi s}\Big\{(1+s)\,\mathrm{Si}\,[\varepsilon_0(1+s)]-$$
$$-(1-s)\,\mathrm{Si}\,[\varepsilon_0(1-s)]-2\,\frac{\sin\varepsilon_0\times\sin(s\,\varepsilon_0)}{\varepsilon_0}\Big\}+$$
$$+\frac{1}{\pi^2 s}\Big\{(1+s)\,\mathrm{Si}^2[\varepsilon_0(1+s)]+$$
$$+2\,\frac{\cos[\varepsilon_0(1+s)]\times\mathrm{Si}[\varepsilon_0(1+s)]}{\varepsilon_0}-\frac{\mathrm{Si}[2\varepsilon_0(1+s)]}{\varepsilon_0}-$$
$$-(1-s)\,\mathrm{Si}^2[\varepsilon_0(1-s)]-$$
$$2\,\frac{\cos[\varepsilon_0(1-s)]\times\mathrm{Si}[\varepsilon_0(1-s)]}{\varepsilon_0}+\frac{\mathrm{Si}[2\varepsilon_0(1-s)]}{\varepsilon_0}\Big\}+$$
$$+\frac{2}{\pi^2}\int_0^{+1}\mathrm{Si}\,[\varepsilon_0(1+sy)\times\mathrm{Si}\,[\varepsilon_0(1-sy)]\,dy.$$

Pour calculer $I(\infty)$, revenons à la formule initiale (12) où, à l'aide de (14), on peut écrire

$$\Big|\Phi(\infty,sy)\Big|^2=$$
$$=\Big|\sqrt{2\pi}\,f(-sy)-\sqrt{\frac{2}{\pi}}\int_{-\infty}^{+\infty}\frac{\sin\varphi}{\varphi}\,e^{-isy\varphi}\,d\varphi\Big|^2=$$
$$=2\pi\Big|f(-sy)\Big|^2.$$

Finalement, on obtient

$$I(\infty)=\frac{1}{4\pi\gamma}\int_{-\infty}^{+\infty}\gamma(y)\times2\pi\Big|f(-sy)\Big|^2dy=1.$$

Le contraste s'écrit

$$(31)\quad \underset{s<1}{C(\varepsilon_0,s)}=\frac{2}{\pi s}\Big\{(1+s)\,\mathrm{Si}\,[\varepsilon_0(1+s)]-$$
$$-(1-s)\,\mathrm{Si}\,[\varepsilon_0(1-s)]-2\,\frac{\sin\varepsilon_0\times\sin s\,\varepsilon_0}{\varepsilon_0}\Big\}$$

$$\frac{1}{\pi^2 s}\Big\{(1+s)\,\mathrm{Si}^2[\varepsilon_0(1+s)]+$$
$$+2\,\frac{\cos[\varepsilon_0(1+s)]\times\mathrm{Si}[\varepsilon_0(1+s)]}{\varepsilon_0}-$$
$$-\frac{\mathrm{Si}[2\varepsilon_0(1+s)]}{\varepsilon_0}-(1-s)\,\mathrm{Si}^2[\varepsilon_0(1-s)]$$
$$-2\,\frac{\cos[\varepsilon_0(1-s)]\times\mathrm{Si}[\varepsilon_0(1-s)]}{\varepsilon_0}+$$
$$+\frac{\mathrm{Si}[2\varepsilon_0(1-s)]}{\varepsilon_0}\Big\}-$$
$$-\frac{2}{\pi^2 s}\int_0^{+s}\mathrm{Si}[\varepsilon_0(1+y)]\times\mathrm{Si}[\varepsilon_0(1-y)]\,dy\,.$$

En faisant $s=0$ dans cette formule, on obtient le contraste d'un ruban opaque éclairé en lumière cohérente.

$$(32)\quad C_{s<1}(\varepsilon_0, s=0)=C_c(\varepsilon_0)=\frac{4}{\pi}\,\mathrm{Si}\,\varepsilon_0-\frac{4}{\pi^2}\,\mathrm{Si}^2\,\varepsilon_0\,.$$

Dans le cas particulier d'un ruban très étroit et d'un instrument de faible ouverture (tel que ε_0^2 soit négligeable par rapport à ε_0), on trouve

$$(33)\qquad C_c=\frac{4}{\pi}\,\varepsilon_0=\frac{4}{\pi}\,k\,n_0\sin\alpha_0\,\eta.$$

Si, de plus, on suppose $n_0=1$ et $\sin\alpha_0=\alpha_0$, on retrouve le résultat de Maréchal [7] :

$$(34)\qquad C_c=\frac{4\pi}{\lambda}\,\alpha_0\times 2\,\eta,$$

où 2η est la largeur du ruban opaque.

b) $s>1$. — En procédant de la même façon que précédemment, on obtient

$$(35)\quad C(\varepsilon_0, s)_{s>1}=\frac{4}{\pi}\Big\{\mathrm{Si}\,2\varepsilon_0-\frac{\sin^2\varepsilon_0}{\varepsilon_0}\Big\}-$$
$$-\frac{1}{\pi^2}\Big\{(1+s)\,\mathrm{Si}^2[\varepsilon_0(1+s)]-(1-s)\,\mathrm{Si}^2[\varepsilon_0(1-s)]+$$
$$+2\,\frac{\cos[\varepsilon_0(1+s)]\,\mathrm{Si}[\varepsilon_0(1+s)]}{\varepsilon_0}-$$
$$-2\,\frac{\cos[\varepsilon_0(1-s)]\,\mathrm{Si}[\varepsilon_0(1-s)]}{\varepsilon_0}-\frac{\mathrm{Si}\,2a(1+s)}{a}+$$
$$+\frac{\mathrm{Si}\,2a(1-s)}{a}\Big\}-\frac{2}{\pi^2}\int_0^{+s}\mathrm{Si}[\varepsilon_0(1+y)]\times$$
$$\times\,\mathrm{Si}[\varepsilon_0(1-y)]\,dy.$$

Le contraste d'un ruban opaque qu'on parviendrait à éclairer en lumière cohérente serait donné par

$$(36)\qquad C_i=\frac{2}{\pi}\Big\{\mathrm{Si}\,2\varepsilon_0-\frac{\sin^2\varepsilon_0}{\varepsilon_0}\Big\}.$$

Dans le cas particulier d'un ruban, tel que ε_0 soit petit, on retrouve le résultat de Maréchal [7] :

$$(37)\qquad C_i=\frac{2\varepsilon_0}{\pi},$$

c'est-à-dire

$$(38)\qquad C_i=\frac{C_c}{2}.$$

En comparant les formules (32) et (36), on voit que, seulement dans le cas d'un objectif peu ouvert et d'un ruban étroit, on peut écrire l'égalité (38).

La figure 5 représente les variations du contraste en fonction du rapport s pour différentes valeurs de $2\varepsilon_0\left(\frac{2\varepsilon_0}{\pi}=2U\right.$ est la largeur géométrique du ruban mesurée en unité optique).

On voit qu'en microscopie, où on utilise couramment des rapports d'ouvertures allant de 0 à 1, le contraste ne varie pratiquement pas pour les faibles valeurs de ε_0. Pour les valeurs de ε_0 dépassant 1,5 le contraste diminue quand le rapport d'ouvertures s'approche de 1.

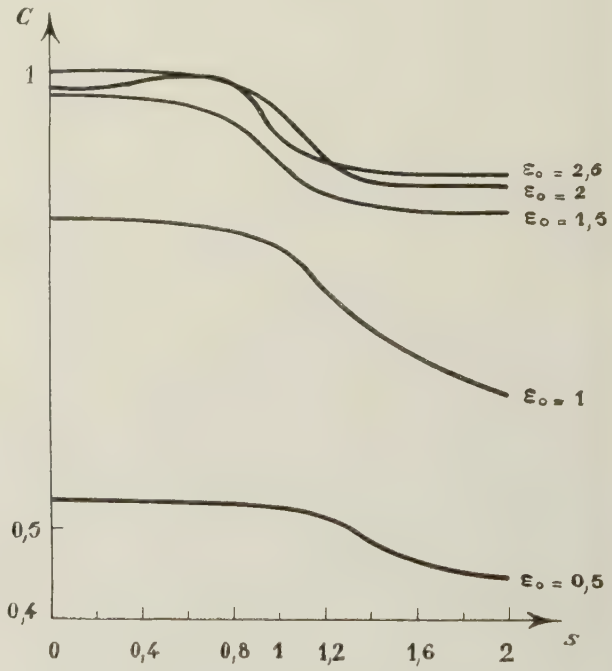

Fig. 5. — Contraste de différents rubans opaques de largeur $2\varepsilon_0$ en fonction du rapport d'ouvertures s.

Sur la figure, les courbes $C=f(s)$ pour $\varepsilon_0=2$ et $\varepsilon_0=2{,}5$ se coupent. Pour les valeurs assez grandes de ε_0, la définition qu'on a choisie pour le contraste ne convient plus : l'éclairement au centre devient plus grand que l'éclairement en un point voisin du centre.

Pour s allant de 1 à l'infini, le contraste diminue toujours.

C. Image d'un demi-plan. (Fig. 6). — Dans ce cas, on a

$$O(v)=\begin{cases}0 \text{ pour } -\infty<v<0\\ 1 \text{ pour } 0<v<+\infty.\end{cases}$$

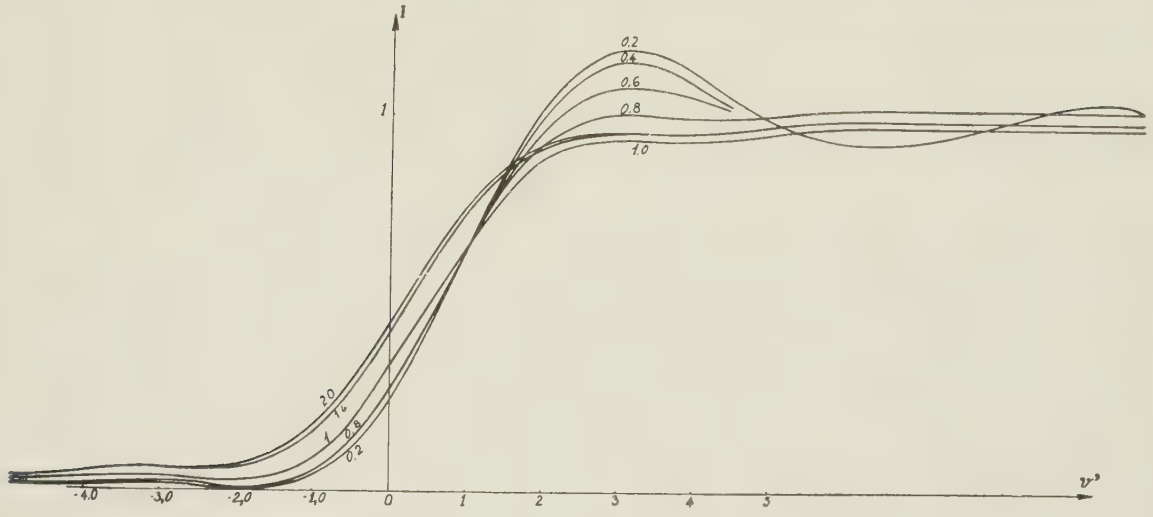

Fig. 6. — Image d'un bord de plage en fonction du rapport d'ouvertures s.

La formule (8) devient

$$\Phi(\nu',sy) = \int_{\mathfrak{n}}^{+\infty} F(\nu'-\nu)\, e^{isy\nu}\, d\nu =$$

$$= e^{isy\nu'} \int_{-\infty}^{\nu'} F(w)\, e^{-isyw}\, dw$$

$$= e^{isy\nu'} \sqrt{\frac{2}{\pi}} \left\{ \int_{-\infty}^{+\nu'} \frac{\sin w}{w} \cos syw\, dw - \right.$$

$$\left. - i \int_{-\infty}^{+\nu'} \frac{\sin w}{w} \sin syw\, dw \right.$$

$$= \sqrt{\frac{2}{\pi}}\, e^{isy\nu'} \left\{ \frac{\pi}{2}\, \varphi(sy) + \frac{1}{2} \left[\mathrm{Si}[\nu'(1+sy)] + \right.\right.$$

$$\left.\left. + \mathrm{Si}[\nu'(1-sy)] \right] + \frac{i}{2} \left[\mathrm{Ci}[\nu'(1+sy)] - \right.\right.$$

$$\left.\left. - \mathrm{Ci}[\nu'(1-sy)] \right] \right\}$$

où

$$\varphi(sy) = \begin{cases} 1 \text{ pour } |sy| < 1 \\ 0 \text{ pour } |sy| > 1. \end{cases}$$

La formule générale de l'éclairement est alors

$$I(\nu') = \frac{A}{2\pi} \int_{-\infty}^{+\infty} \gamma(y) \left\{ [\pi.\varphi(sy) + H]^2 + K^2 \right\} dy$$

où

$$\begin{cases} H = \mathrm{Si}[\nu'(1+sy)] + \mathrm{Si}[\nu'(1-sy)] \\ K = \mathrm{Ci}(\nu'(1+sy)] - \mathrm{Ci}[\nu'(1-sy)]. \end{cases}$$

Comme précédemment, il faut distinguer le cas $s < 1$ du cas $s > 1$.

a) $s < 1$. — L'éclairement dans le plan image, en l'absence de l'objet, est

$$4\pi A\gamma.$$

Compte tenu des définitions de $\varphi(sy)$ et de $\gamma(y)$, on a

$$39)\quad I(\nu') = \frac{1}{8\pi^2} \left\{ 2\pi^2 + 2\pi \int_{-1}^{+1} H\, dy + \right.$$

$$\left. + \int_{-1}^{+1} H^2\, dy + \int_{-1}^{+1} K^2\, dy \right\}$$

$$= + \frac{1}{4} + \frac{1}{2\pi s} \left\{ (1+s)\,\mathrm{Si}[\nu'(1+s)] - \right.$$

$$\left. - (1-s)\,\mathrm{Si}[\nu'(1-s)] - 2\,\frac{\sin\nu' \times \sin s\nu'}{\nu'} \right\}$$

$$+ \frac{1}{4\pi^2 s} \left\{ (1+s)\,\mathrm{Si}^2[\nu'(1+s)] - (1-s)\,\mathrm{Si}^2[\nu'(1-s)] + \right.$$

$$+ \frac{2\cos[\nu'(1+s)] \times \mathrm{Si}[\nu'(1+s)]}{\nu'} -$$

$$- 2\,\frac{\cos[\nu'(1-s)] \times \mathrm{Si}[\nu'(1-s)]}{\nu'} + (1+s)\,\mathrm{Ci}^2(\nu'(1+s)] -$$

$$- (1-s)\mathrm{Ci}^2[\nu'(1-s)] - 2\,\frac{\sin[\nu'(1+s)] \times \mathrm{Ci}[\nu'(1+s)]}{\nu'} +$$

$$\left. + 2\,\frac{\sin\nu'(1-s) \times \mathrm{Ci}\,\nu'(1-s)}{\nu'} \right\} +$$

$$+ \frac{1}{2\pi^2} \left\{ \int_0^{+1} \mathrm{Si}[\nu'(1+sy)] \times \mathrm{Si}[\nu'(1-sy)]\, dy - \right.$$

$$\left. - \int_0^{+1} \mathrm{Ci}[\nu'(1+sy)] \times \mathrm{Ci}[\nu'(1-sy)]\, dy \right\}.$$

A la limite, pour $s = 0$, on retrouve l'expression bien connue de l'éclairement dans l'image d'un demi-plan éclairé en lumière cohérente.

$$(40)\qquad I_c = \left\{ \frac{1}{2} + \frac{\mathrm{Si}\,\nu'}{\pi} \right\}^2.$$

Formons les dérivées première et seconde de cette équation,

$$(41)\quad \frac{\partial I_c}{\partial \varphi'} = \frac{1}{\pi}\frac{\sin\varphi'}{\varphi'}\left\{1+\frac{2}{\pi}\,\mathrm{Si}\,\varphi'\right\},$$

$$\frac{\partial^2 I_c}{\partial \varphi'^2} = \frac{1}{\pi}\left\{\left(\frac{\cos\varphi'}{\varphi'}-\frac{\sin\varphi'}{\varphi'^2}\right)\left(1+\frac{2}{\pi}\,\mathrm{Si}\,\varphi'\right)+\right.$$
$$\left.+\frac{2}{\pi}\left(\frac{\sin\varphi'}{\varphi'}\right)^2\right\}.$$

Pour toutes les valeurs positives de φ', les extrema de I_c sont donnés par les zéros de $\frac{\sin\varphi'}{\varphi'}$; pour les $\varphi'<0$, les zéros de $\frac{\sin\varphi'}{\varphi'}$ donnent des maxima $|\varphi'| = k\pi$ et les zéros de $\left(1+\frac{2}{\pi}\,\mathrm{Si}\,\varphi'\right)$ fournissent des minima. On retrouve le résultat bien connu que les franges sont deux fois plus serrées dans la partie sombre que dans la partie claire de l'image.

Dans le cas général, la dérivée de l'éclairement par rapport à φ' s'écrit

$$(42)\quad \frac{\partial I(\varphi')}{\partial \varphi'} = -\frac{1}{2\pi^2}\frac{\sin\varphi'}{\varphi'}\int_{-1}^{+1}[\cos(s\,y\,\varphi')\times$$
$$\times(\pi\varphi+H)-K\sin s\,y\,\varphi']\,dy\,.$$

Le facteur $\frac{\sin\varphi'}{\varphi'}$ est mis en évidence, comme dans le cas cohérent (41).

Alors, il y a au moins autant de franges qu'en éclairage cohérent.

Si, dans (42), on fait $\varphi' = 0$, on voit que la fonction est indépendante de s : toutes les courbes représentant l'intensité ont même pente au point $\varphi' = 0$ quel que soit le rapport des ouvertures. Si on explicite (42), on a

$$(43)\quad \frac{\partial I(\varphi')}{\partial \varphi'} =$$
$$= \frac{1}{\pi^2}\frac{\sin\varphi'}{\varphi'}\left\{\frac{\sin\varphi' s}{\varphi' s}\left[\pi+\mathrm{Si}\,[\varphi'(1+s)]+\right.\right.$$
$$\left.+\,\mathrm{Si}\,[\varphi'(1-s)]\right]$$
$$+\frac{\cos\varphi' s}{\varphi' s}\left[\mathrm{Ci}\,[\varphi'(1+s)]-\mathrm{Ci}\,[\varphi'(1-s)]\right]-$$
$$\left.-\frac{\cos\varphi'}{\varphi' s}\log\frac{1+s}{1-s}\right\}.$$

Pour les petites valeurs de s telles que s^2 soit négligeable par rapport à s, cette formule devient

$$\underset{s\ \text{petit}}{\frac{\partial I(\varphi')}{\partial \varphi'}} = -\frac{2}{\pi^2}\frac{\sin\varphi'}{\varphi'}\times\frac{\sin\frac{\varphi' s}{2}}{\frac{\varphi' s}{2}}\times$$
$$\times\left\{\left(\frac{\pi}{2}+\mathrm{Si}\,\varphi'\right)\cos\frac{\varphi' s}{2}-s\cos\varphi'\sin\frac{\varphi' s}{2}\right\}.$$

Cette équation s'annule, en particulier, pour les zéros du crochet. On voit que, dans la partie sombre de l'image, la première frange noire s'écarte plus du bord géométrique que la première frange noire de l'image cohérente. Si $s = 1$, (43) et la dérivée de (43) donnent

$$\underset{s\to 1}{\frac{\partial I(\varphi')}{\partial \varphi'}} = \frac{1}{\pi^2}\frac{\sin\varphi'}{\varphi'}\left\{\frac{\sin\varphi'}{\varphi'}\left[\pi+\mathrm{Si}\,2\varphi'\right]+\right.$$
$$\left.+\frac{\cos\varphi'}{\varphi'}\left[\mathrm{Ci}\,2\varphi'-\log 2\gamma\varphi'\right]\right\}$$

$$\frac{\partial^2 I}{\partial \varphi'^2} = \frac{1}{\pi^2}\left(\frac{\cos\varphi'}{\varphi'}-\frac{\sin\varphi'}{\varphi'^2}\right)\left\{\frac{\sin\varphi'}{\varphi'}\left[\pi+\mathrm{Si}\,2\varphi'\right]+\right.$$
$$\left.+\frac{\cos\varphi'}{\varphi'}\left[\mathrm{Ci}\,2\varphi'-\log 2\gamma\varphi'\right]\right\}$$
$$+\frac{1}{\pi^2}\frac{\sin\varphi'}{\varphi'}\left\{\left(\frac{\cos\varphi'}{\varphi'}-\frac{\sin\varphi'}{\varphi'^2}\right)\left(\pi+\mathrm{Si}\,2\varphi'\right)+\right.$$
$$+\frac{\sin\varphi'}{\varphi'}\frac{\sin 2\varphi'}{\varphi'}+\frac{\cos\varphi'}{\varphi'}\left(\frac{\cos 2\varphi'}{\varphi'}-\frac{1}{\varphi'}\right)-$$
$$\left.-\left(\frac{\sin\varphi'}{\varphi'}+\frac{\cos\varphi'}{\varphi'^2}\right)\left(\mathrm{Ci}\,2\varphi'-\log 2\gamma\varphi'\right)\right\}.$$

Pour $\varphi' = n\pi$ (ces valeurs de φ' correspondent au zéro du cas cohérent), la dérivée seconde est tantôt négative, tantôt positive.

Dans le cas où le rapport des ouvertures est égal à 1, l'image présente encore des franges.

b) $s \geqslant 1$. — Le facteur de normalisation est $\frac{4\pi A\gamma}{s}$.

On a

$$I(\varphi') = \frac{s}{2\pi^2}\int_{-1/s}^{+1/s}(\pi^2\varphi^2+2\pi\varphi H)\,dy+$$
$$+\frac{s}{8\pi^2}\int_{-1}^{+1}(H^2+K^2)\,dy.$$

$$(44)\quad I(\varphi') = \frac{1}{4}+\frac{1}{\pi}\left[\mathrm{Si}\,2\varphi'-\frac{\sin^2\varphi'}{\varphi'}\right]+$$
$$+\frac{1}{4\pi^2}\left\{(1+s)\,\mathrm{Si}^2[\varphi'(1+s)]-\right.$$
$$-(1-s)\,\mathrm{Si}^2[\varphi'(1-s)]+2\frac{\cos[\varphi'(1+s)]\,\mathrm{Si}[\varphi'(1+s)]}{\varphi'}-$$
$$-2\frac{\cos[\varphi'(1-s)]\,\mathrm{Si}[\varphi'(1-s)]}{\varphi'}+(1+s)\,\mathrm{Ci}^2[\varphi'(1+s)]-$$
$$-2\frac{\sin[\varphi'(1+s)]\,\mathrm{Ci}[\varphi'(1+s)]}{\varphi'}+$$
$$+2\frac{\sin[\varphi'(1-s)]\,\mathrm{Ci}[\varphi'(1-s)]}{\varphi'}-$$
$$\left.-(1-s)\,\mathrm{Ci}^2[\varphi'(1-s)]\right\}+\frac{s}{2\pi^2}\int_0^{+1}\mathrm{Si}[\varphi'(1+sy)]\times$$
$$\times\,\mathrm{Si}[\varphi'(1-sy)]\,dy-\int_0^{+1}\mathrm{Ci}\,[\varphi'(1+sy)]\times$$
$$\times\,\mathrm{Ci}\,[\varphi'(1-sy)]\,dy.$$

En prenant la limite, pour $s\to\infty$, on obtient l'éclairement en lumière incohérente :

$$(45)\quad I(\varphi') = \frac{1}{2} + \frac{1}{\pi}\left[\mathrm{Si}\,2\varphi' - \frac{\sin^2\varphi'}{\varphi'}\right].$$

4. Optique et objet à symétrie circulaire.

A. ***Image d'un disque opaque.*** — L'équation générale de l'éclairement en un point P′ de l'image est

$$(46)\quad \mathfrak{I}(\mathrm{P}') = A\int_{-\infty}^{+\infty}\gamma(\mathrm{M})\left|\int_{-\infty}^{+\infty} O(\mathrm{P}).F(\mathrm{P}'-\mathrm{P})\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{P}\right|^2 d\mathrm{M}$$

où A est le facteur de normalisation qui permet de rendre égal à l'unité l'éclairement de l'image en l'absence de l'objet.

Comme précédemment, il faut distinguer le cas $s < 1$ et le cas $s > 1$

$$s < 1\ ,\qquad A = \frac{1}{4\pi^3\gamma}\ ;$$

$$s > 1\ ,\qquad A = \frac{s^2}{4\pi^3\gamma}\ .$$

Si l'objet est un disque opaque de rayon égal à ε_0, on a

$$O(\mathrm{P}) = \begin{cases} 0 \text{ pour } |\mathrm{P}| < \varepsilon_0 \\ 1 \text{ pour } |\mathrm{P}| > \varepsilon_0\ . \end{cases}$$

La fonction Φ peut s'écrire

$$\Phi(\mathrm{P}', \mathrm{M}, s) = \int_{-\infty}^{+\infty} F(\mathrm{P}'-\mathrm{P})\,\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{P} - \int_{\text{Disque}} F(\mathrm{P}'-\mathrm{P})\,\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{P} = 2\pi f(s\,\mathrm{M}) - \int_{\text{Disque}} F(\mathrm{P}'-\mathrm{P})\,\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{P}\ .$$

Ce qui donne, pour l'éclairement,

$$(47)\quad \mathfrak{I}(\mathrm{P}') = A\left\{4\pi^2\int_{-\infty}^{+\infty}\gamma(\mathrm{M})\,f^2(s\,\mathrm{M})\,d\mathrm{M} - 4\pi\int_{-\infty}^{+\infty}\gamma(\mathrm{M})f(s\,\mathrm{M})\,\mathfrak{R}\int_{\text{Disque}} F(\mathrm{P}'-\mathrm{P})\,\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{P} + \int_{-\infty}^{+\infty}\gamma(\mathrm{M})\left|\int_{\mathrm{D}} F(\mathrm{P}'-\mathrm{P})\,\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{P}\right|^2 d\mathrm{M}\right\}.$$

a) $s < 1$. — La transformée de Fourier de $\gamma(\mathrm{M})$ est

$$\Gamma(s\,\mathrm{P}) = \frac{1}{2\pi}\int_0^{+1}\gamma(\mathrm{M})\,\mathbf{e}^{-is\mathrm{MP}}\,d\mathrm{M} = \gamma\,\frac{\mathfrak{J}_1(s\,\rho)}{s\,\rho}\ ,$$

avec $\gamma(\mathrm{M})$ $\begin{cases} \gamma \text{ pour } |\mathrm{M}| < 1 \\ 0 \text{ pour } |\mathrm{M}| > 1\ . \end{cases}$

La valeur de la fonction F est

$$F(\mathrm{P}) = \frac{1}{2\pi}\int f(\mathrm{M})\,\mathbf{e}^{i\mathrm{MP}}\,d\mathrm{M} \qquad \frac{\mathfrak{J}_1(\rho)}{\rho}$$

car $f(\mathrm{M}) = \begin{cases} 1 \text{ pour } |\mathrm{M}| < 1 \\ 0 \text{ pour } |\mathrm{M}| > 1\ . \end{cases}$

L'équation donnant l'éclairement prend la forme

$$(48)\quad I(\mathrm{P}') = 1 - \frac{2}{\pi\gamma_s}\int_{\mathrm{D}} F(\mathrm{P}'-\mathrm{P})\,\mathrm{T}(s\,\mathrm{P})\,d\mathrm{P} + \frac{1}{2\pi^2\gamma_s}\int_{\mathrm{D}}\int_{\mathrm{D}} F(\mathrm{P}'-\mathrm{P}_1)\,F^*(\mathrm{P}'-\mathrm{P}_2)\times \Gamma[s(\mathrm{P}_1-\mathrm{P}_2)]\,d\mathrm{P}_1\,d\mathrm{P}_2$$

$$= 1 - \frac{2}{\pi}\int_{\mathrm{D}}\frac{\mathfrak{J}_1(\mathrm{P}'-\mathrm{P}_1)}{\mathrm{P}'-\mathrm{P}_1}\times\frac{\mathfrak{J}_1(s\,\mathrm{P})}{s\,\mathrm{P}}\,d\mathrm{P} + \frac{1}{2\pi^2}\int_{\mathrm{D}}\int_{\mathrm{D}}\frac{\mathfrak{J}_1(\mathrm{P}'-\mathrm{P}_1)}{\mathrm{P}'-\mathrm{P}_1}\times\frac{\mathfrak{J}_1(\mathrm{P}'-\mathrm{P}_2)}{\mathrm{P}'-\mathrm{P}_2}\times \times\frac{\mathfrak{J}_1[s(\mathrm{P}_1-\mathrm{P}_2)]}{s(\mathrm{P}_1-\mathrm{P}_2)}\,d\mathrm{P}_1\,d\mathrm{P}_2.$$

La résolution de cette équation donne une expression très compliquée qu'il est difficile de discuter, toutefois, dans le cas d'un disque et d'une ouverture numérique de l'objectif tels qu'on puisse négliger ε_0^4 par rapport à ε_0^2, on obtient

$$(49)\qquad I(\mathrm{P}') = 1 - \varepsilon_0^2\,\frac{\mathfrak{J}_1(\rho')}{\rho'}\ .$$

Eclairement de l'image d'un disque opaque éclairé en lumière cohérente. — En faisant $s = 0$ dans la formule (48), on obtient tout de suite l'image du disque noir en lumière cohérente.

$\dfrac{\mathfrak{J}_1[s(\mathrm{P}_1 \quad \mathrm{P}_2)]}{s(\mathrm{P}_1-\mathrm{P}_2)}$ prend la valeur 1/2 pour $s = 0$.

Les variables se séparent et on a

$$(50)\quad I_c(\mathrm{P}') = I_c(\rho') = \left\{1 - \frac{1}{2\pi}\int_{\mathrm{D}}\frac{\mathfrak{J}_1(\mathrm{P}'-\mathrm{P})}{\mathrm{P}'-\mathrm{P}}\,d\mathrm{P}\right\}^2$$

$$= \left\{1 - \frac{1}{2\pi}\int_0^{2\pi}\int_0^{+\varepsilon_0}\sum_{m=1,2,3,\ldots}^{\infty} 2m\,\frac{\mathfrak{J}_m(\rho')}{\rho'}\,\frac{\mathfrak{J}_m(\rho)}{\rho}\times \times\frac{\sin m\varphi}{\sin\varphi}\,\rho\,d\rho\,d\varphi\right\}^2$$

$$= \left\{1 - 2\int_0^{+\varepsilon_0}\sum_{m=1,3,5}^{\infty} m\,\frac{J_m(\rho')}{\rho'}\,\mathfrak{J}_m(\rho)\,d\rho\right\}^2$$

$$= \left\{1 - 4\sum_{m=1,3,5,\ldots}^{\infty} m\,\frac{\mathfrak{J}_m(\rho')}{\rho'}\sum_{p=0}^{\infty}\mathfrak{J}_{m+2p+1}(\varepsilon_0)\right\}^2\ .$$

Dans le cas où ε_0 est suffisamment petit pour qu'on puisse négliger ε_0^4 par rapport à ε_0^2, on a

$$(51)\qquad I_c(\rho') = 1 - \varepsilon_0^2\,\frac{\mathfrak{J}_1(\rho')}{\rho'} = 1 - \frac{\mathfrak{J}_1(\rho')}{\rho'}\times\frac{4\pi\sin^2\alpha_0}{\lambda^2}\,\pi\,\xi_0^2\ ,$$

où $\pi\,\xi_0^2$ est la surface du petit disque.

En comparant ce résultat avec la formule (49), on voit que l'ouverture numérique du condenseur n'agit pas sur l'image d'un disque petit, autrement dit, un petit objet microscopique regardé avec une faible ouverture est toujours vu comme un objet cohérent.

Lorsque $\rho' = 0$ dans (50), on retrouve la formule connue donnant l'éclairement au centre de l'image cohérente :

$$I_c(0) = J_0^2(\varepsilon_0). \tag{52}$$

b) $s > 1$. — La transformée de $f(s\,M)$ est maintenant

$$\frac{1}{2\pi}\int_0^{2\pi}\int_0^{1/s} e^{-isMP}\,dM = \frac{1}{s^2}\frac{J_1(P)}{P}.$$

A partir de la formule (47), on obtient immédiatement

$$I(P') = 1 - \frac{2}{\pi}\int_D \frac{J_1(P'-P)}{P'-P}\cdot\frac{J_1(P)}{P}\,dP + \frac{s^2}{2\pi^2}\int_D\int_D \frac{J_1(P'-P_1)}{P'-P_1}\cdot\frac{J_1(P'-P_2)}{P'-P_2}\,\frac{J_1 s(P_1-P_2)}{s(P_1-P_2)}\,dP_1\,dP_2$$

Cette formule se différencie de la formule (48) par le 2e terme qui n'est plus fonction de s et par le 3e terme qui est le même que dans le cas précédent au facteur s^2 près.

Image d'un disque opaque sur fond incohérent. — L'image d'un disque opaque sur fond incohérent a été calculée récemment par WEINSTEIN [9]. C'est pour cette raison que nous ne la donnons pas ici. Elle peut aussi être obtenue en faisant la limite, pour s tendant vers l'infini, de l'équation (46) normalisée pour $s > 1$.

L'éclairement au centre de l'image incohérente est

$$I_i(o) = J_0^2(\varepsilon_0) + J_1^2(\varepsilon_0). \tag{53}$$

B. *Contraste de l'image d'un disque opaque*

a) $s < 1$. — L'éclairement au centre de l'image est, à partir de (48),

$$I(0) = 1 - \frac{2}{\pi}\int_D \frac{J_1(P)}{P}\cdot\frac{J_1(sP)}{sP}\,dP + \frac{1}{2\pi^2}\int_D\int_D \frac{J_1(P_1)}{P_1}\frac{J_1(P_2)}{P_2}\times\frac{J_1[s(P_1-P_2)]}{s(P_1-P_2)}\,dP_1\,dP_2$$

$$= 1 - \frac{4}{s}\int_0^{+\varepsilon_0}\frac{J_1(\rho)\,J_1(s\rho)}{\rho}\,d\rho + \frac{4}{s^2}\sum_{\mu=1,3,5\ldots}^{\infty}\mu\left\{\int_0^{+\varepsilon_0}\frac{J_\mu(s\rho)\,J_1(\rho)}{\rho}\,d\rho\right\}^2.$$

L'éclairement, loin du disque, est évidemment $I(\infty) = 1$, en raison de la normalisation. Le contraste (5) peut s'écrire de façon condensée à l'aide des fonctions hypergéométriques d'ordre 2 :

$$C = \frac{4}{s^2}\sum_{m=0}^{\infty}(-1)^m\left(\frac{1}{2}s\varepsilon_0\right)^{2m+2}\frac{{}_2F_1(-m,-1-m;2;1/s^2)}{(2m+2)\,m!\;(m+1)!} - \frac{4}{s^4}\sum_{m=1,3,5\ldots}^{\infty}\mu\left\{\sum_{n=0}^{\infty}\frac{(-1)^n\left(\frac{1}{2}s\varepsilon_0\right)^{2n+\mu+1}{}_2F_1(-n,-\mu-n;2;1/s^2)}{(2n+\mu+1)n!\;\Gamma(\mu+n+1)}\right\}^2$$

ou, encore, avec les fonctions de BESSEL,

$$C = 1 - 2\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty} J_n(\varepsilon_0)J_m(\varepsilon_0)s^{n+m}\times\frac{(-1)^\nu}{\nu!\,\Gamma(n+\nu+1)}\frac{(\varepsilon_0 s)^{n+2\nu}}{(2)^{n+2\nu}}\times\frac{(-1)^\mu}{\mu!\,\Gamma(\mu+m+1)}\times\left(\frac{\varepsilon_0 s}{2}\right)^{m+2\mu}\times\frac{1}{2(n+m+\mu+\nu+1)}.$$

A la limite, lorsque s tend vers zéro, on trouve le contraste du point éclairé en lumière cohérente

$$C_0 = 1 - J_0^2(\varepsilon_0). \tag{56}$$

Dans le cas particulier où on admet que ε_0^4 est négligeable par rapport à ε_0^2, on trouve

$$C_c = \frac{\varepsilon_0^2}{2} = \frac{2\pi\sin^2\alpha_0}{\lambda^2}\pi\,\xi_0^2 = \frac{2\pi\sin^2\alpha_0}{\lambda^2}S. \tag{57}$$

De plus, si $\sin^2\alpha_0 = \alpha_0^2$, on retrouve l'expression de MARÉCHAL [7], S étant la surface du disque.

b) $s > 1$. — Dans ce cas, l'éclairement au centre de l'image est

$$I(0) = 1 - \frac{2}{\pi}\int_D\left[\frac{J_1(P)}{P}\right]^2 dP + \frac{s^2}{2\pi^2}\int_D\int_D \frac{J_1(P'-P_1)}{P'-P_1}\frac{J_1(P'-P_2)}{P'-P_2}\frac{J_1[s(P_1-P_2)]}{s(P_1-P_2)}\,dP_1\,dP_2.$$

Par des calculs analogues aux précédents, on établit le contraste

$$C = 2\left\{1 - J_0^2(\varepsilon_0) - J_1^2(\varepsilon_0)\right\} - \frac{2}{s^2}\sum_{\mu=1,3,5,\ldots}^{\infty}\mu\left\{\sum_{n=0}^{\infty}\frac{(-1)^n\left(\frac{1}{2}s\varepsilon_0\right)^{2n+\mu+1}{}_2F_1(-n,-\mu-n;2;1/s^2)}{2n+\mu+1)\,n!\,\Gamma(\mu+n+1)}\right\}^2. \tag{58}$$

Dans le cas d'un éclairage incohérent, le contraste du point opaque est

$$C_i = 1 - J_0^2(\varepsilon_0) - J_1^2(\varepsilon_0). \tag{59}$$

En faisant la même approximation que dans (58), on retrouve la formule de MARÉCHAL :

$$C_i = \frac{\varepsilon_0^2}{4}.$$

Remarque. — En faisant la différence des équations (57) et (59), on voit que

$$C_c - C_i = \mathfrak{J}_1^2(\varepsilon_0). \tag{60}$$

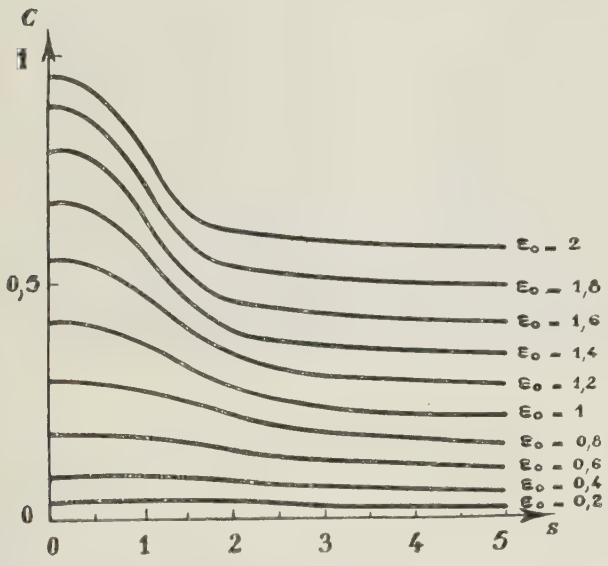

Fig. 7. — Contraste de l'image de différents disques de diamètre $2\varepsilon_0$ en fonction de s.

On remarque que la différence des contrastes cohérent et incohérent augmente avec le diamètre du disque, il est maximum lorsque le disque a le diamètre du 1er anneau de la tache de diffraction.

Dans le cas de l'approximation de Maréchal, cette différence devient

$$C_c - C_i = \frac{\varepsilon_0^2}{4} = \frac{\pi \sin^2 \alpha_0}{\lambda^2} \pi \xi_0^2 = \pi \frac{\sin^2 \alpha_0}{\lambda^2} S. \tag{61}$$

La figure 7 donne l'allure du contraste, pour différentes valeurs du rayon du disque, en fonction du rapport des ouvertures.

5. **Etude expérimentale.** — Toutes les photomicrographies ont été faites à l'aide d'un microscope équipé d'un condenseur panchratique. Nous avons choisi un objectif bien corrigé de l'aberration sphérique et d'ouverture faible pour rester dans le domaine valable de la théorie des ondes scalaires ($g = 13 \times$; $n_0 \sin \alpha_0 = 0{,}20$). Nous avons utilisé un oculaire de projection de grandissement $6 \times$. Le grandissement définitif est de l'ordre de 1 000. Les négatifs ont été obtenus sur microfilm Lumière développé à l'Isonal. Nous avons photographié les divers objets dont nous avons calculé théoriquement l'image. La difficulté était d'obtenir des tests à la fois opaques et de dimensions microscopiques convenables.

A. Le fil opaque. — Nous avons utilisé un fil d'araignée aluminié par projection sous vide. Son diamètre, légèrement inférieur au micron, était plus petit que le diamètre du premier anneau de la tache de diffraction de l'objectif. Nous avons photographié ce fil avec différents rapports d'ouvertures : $s = 0$; 0,25 ; 0,50 ; 1,50 (fig. 8).

B. ***Le bord de plage.*** — Le bord de plage a été obtenu par une trace d'encre de Chine sur une lame porte-objet. Il a été photographié pour $s = 0$; 0,25 ; 1,0 ; 1,50 (fig. 9). On voit que, pour un rapport d'ouverture égal à 1, les franges mises en évidence théoriquement sont trop faibles pour être visibles sur la photographie.

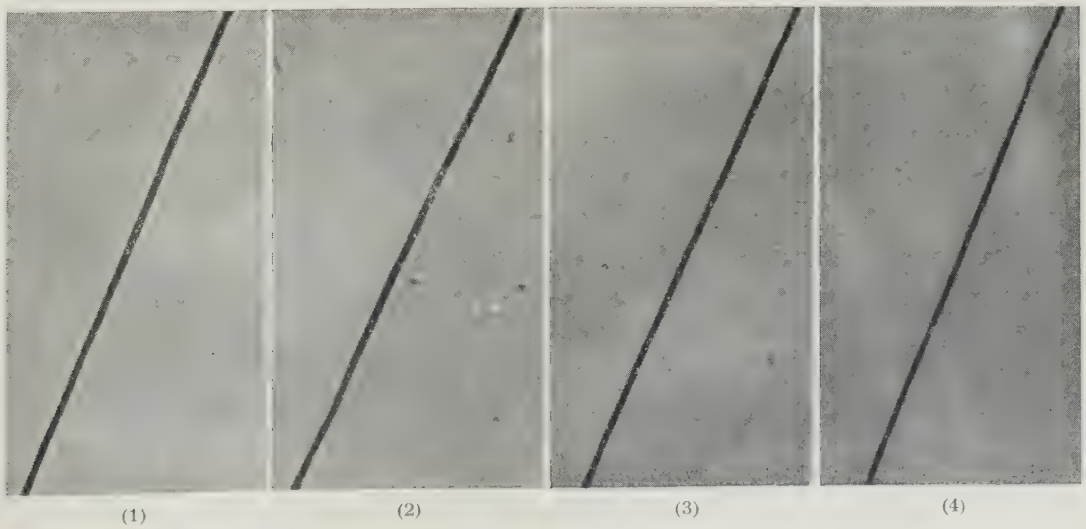

Fig. 8. — Image d'un fil opaque.
(1) $s = 0$; (2) $s = 0{,}25$; (3) $s = 0{,}50$; (4) $s = 1{,}50$.

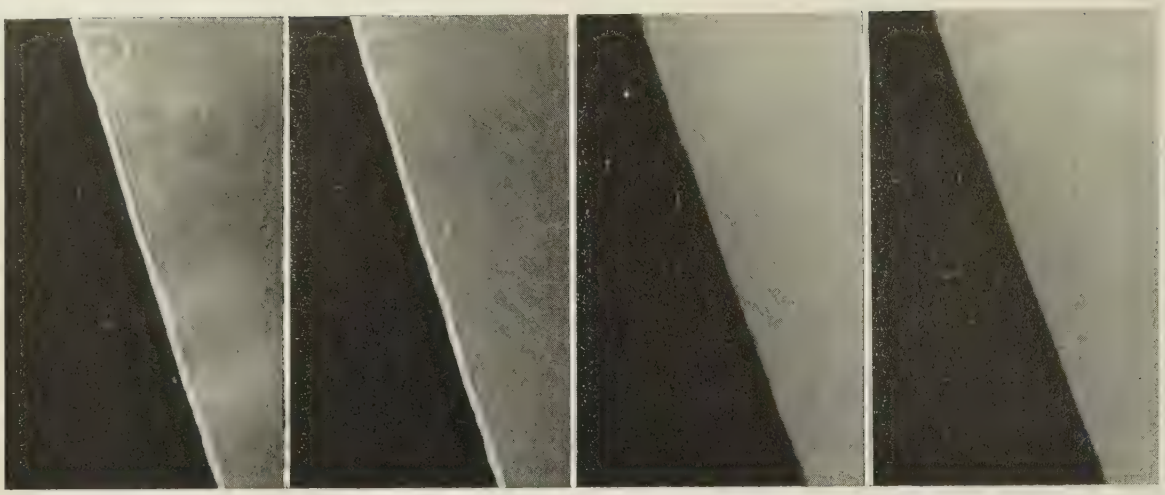

Fig. 9. — Image du bord de plage.
(1) $s = 0$; (2) $s = 0{,}25$; (3) $s = 1{,}0$; (4) $s = 1{,}50$.

C. ***Le point noir.*** — Nous avons pris des staphylocoques très fortement colorés afin de les rendre opaques. Leur diamètre, de l'ordre de 0,5 μ, convenait très bien. Nous donnons des photographies de ces points noirs pour $s = 0$; 0,25 ; 0,50 ; 1,50 (fig. 10).

Tous les résultats expérimentaux sont en bon accord avec la théorie.

Conclusion. — A l'aide des formules développées par Hopkins [2, 3], nous avons calculé les images (et, dans certains cas, le contraste) de 1° un ruban opaque, 2°) un demi-plan et 3°) un disque opaque, en fonction du rapport s des ouvertures du condenseur et de l'objectif.

Dans chaque cas, nous trouvons comme limite la cohérence et l'incohérence. Ceci est possible parce que nous normalisons les formules différemment selon que s est plus petit ou plus grand que 1. Il est clair que ce procédé renferme un certain arbitraire qui est uniquement justifié par le besoin de pouvoir comparer les résultats à des formules connues. L'origine de cet état de choses est l'utilisation d'éléments de surface pour définir l'élément d'intensité : en toute rigueur, on devrait partir d'éléments ponctuels et faire l'addition.

En conclusion, on ne peut pas parler de l'image

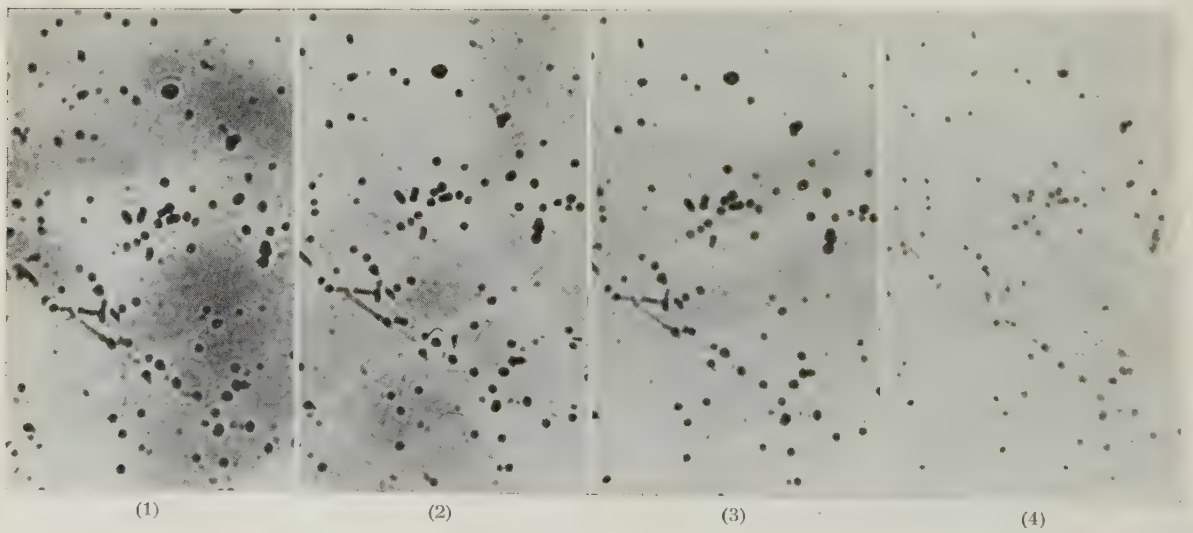

Fig. 10. — Image d'un disque opaque.
(1) $s = 0$; (2) $s = 0{,}25$; (3) $s = 0{,}50$; (4) $s = 1{,}50$.

d'un objet opaque mais « des images » d'un objet opaque : la distribution d'éclairement varie avec le rapport s des ouvertures des systèmes d'éclairage et de reproduction. C'est-à-dire que, si l'on veut caractériser l'objet à partir de l'image, il faut obligatoirement donner le rapport s des ouvertures : il n'y a pas de relation biunivoque entre objet et image.

Qu'il nous soit permis de remercier Mr Maréchal et Mr Nomarski qui ont bien voulu nous faire part de leurs conseils.

Bibliographie

[1] Zernike F., *Physica*, **5**, 1938, 785.
Zernike F., *Proc. Phys. Soc.*, **61**, 1948, 158.
[2] Hopkins H. H., *Proc. Roy. Soc.*, **208**, 1951, 263.
[3] Hopkins H. H., *Proc. Roy. Soc.*, **217**, 1953, 408.
[4] Duffieux P. M., *Rev. d'Opt.*, **32**, 1953, 129.
[5] Blanc-Lapierre A. et Dumontet P. *Rev. Opt.*, **34**, 1955, 1.
Blanc-Lapierre A., Perrot M., Péri G. *Opt. Acta*, **2**, 1955, 1.
[6] Wolf E. *Proc. Roy. Soc.*, **225**, 1954, 96.
[7] Maréchal A. *Rev. Opt.*, **32**, 1953, 649.
[8] Dumontet P. *Publ. Sci. Univ. Alger*, **B1**, 1955, 33.
[9] Weinstein, *J. O. S. A.*, **12**, 1955, 1006.

Manuscrit reçu le 7 septembre 1957.

The influence of scattering within photographic emulsions on resolving power

R. G. Giovanelli

Division of Physics, National Standards Laboratory.
Commonwealth Scientific and Industrial Research Organization, Sydney.

Summary. — *A theoretical discussion is given of the effects on resolving power of the diffusion of radiation within photographic emulsions, particularly for low contrast images.*

Sommaire. — *Influence théorique de la diffusion de l'émulsion photographique sur le pouvoir séparateur, en particulier dans le cas des images à faible contraste.*

Zusammenfassung. — *Der Einfluss der Lichtstreuung innerhalb der photographischen Emulsion auf das Auflösungs vermögen wird theoretisch untersucht, insbesondere bei Abbildungen mit geringem Kontrast.*

I. **Introduction.** — The resolving power of a photographic emulsion is normally measured by illuminating the plate in a series of strips of equal width but alternating brightness, here called the object pattern, and is expressed as the reciprocal of the width of the finest pattern just resolved on the developed plate. It depends on object contrast, expressed here as the ratio of the greater to the lesser illumination in the object pattern, on the image contrast required for resolution, and on several plate factors, of which the most important are probably granularity and the diffusion of light within the emulsion. Quantitative information on the relative importances of the various factors is lacking, however, and it is the purpose of this note to show how the effects of diffusion may be calculated.

The method, which is inevitably very approximate and neglects many features which are of importance in practice, involves evaluating the light distribution within the emulsion when illuminated by a resolving power object pattern, and finding from this the corresponding distribution of the photographic image in the developed plate. The density of the developed plate is obtained by integration throughout the depth of the emulsion, and the variation in density across the surface of the plate gives the contrast in the photographic image. On the assumption that the resolving power is limited by the density variations alone, the resolving power for a given image contrast may be calculated for a given object contrast. These results may be improved by taking account of limits imposed by granularity.

II. **General properties of photographic emulsions.** — Consider an isotropically scattering emulsion of thickness ζ, refractive index of the matrix N, scattering coefficient σ. In addition suppose light is absorbed by the matrix or grains, or both, with absorption coefficient α. The extinction coefficient is then $\varkappa = \sigma + \alpha$, the albedo for single scattering $\varpi_0 = \sigma/\varkappa$ and its complement $\lambda = 1 - \varpi_0$.

Since the diffusion of light within the emulsion depends on ζ, N, $\varkappa$ and ϖ_0, their orders of magnitude in typical emulsions are required.

The refractive index N may be taken as that of gelatine, lying in the range 1.516 — 1.534 according to the Smithsonian Physical Tables. The value assumed in this section is 1.523.

The albedo for single scattering ϖ_0 can be found from the reflectance $\mathfrak{R}$ of a semi-infinite emulsion, prepared effectively by cementing a number of plates together with a liquid whose refractive index is close to that of the emulsion and glass base. The relation between $\mathfrak{R}$ and ϖ_0 can be found from tables [1] of the reflectance of a semi-infinite diffuser of refractive index 1.5 with a perfectly diffusing surface, together with values of $\mathrm{d}\mathfrak{R}/\mathrm{d}N$.

Pitts [2] has shown on the Eddington approximation how for normally incident radiation the diffuse transmittance T and the corresponding diffuse reflec-

tance R of a plane parallel diffusing plate, matrix of refractive index unity, is related to ϖ_0 and to optical thickness $\varkappa\zeta$ (his symbols ε and τ). This relation may be used to derive $\varkappa\zeta$, and hence $\varkappa$, for a layer whose matrix is of refractive index N by noting that the diffuse transmittance T_1 of such a layer may be expressed in terms of its diffuse transmittance T when surrounded by a clear medium of refractive index N (to which Pitts' analysis applies),

$$(2.1) \quad T_1 = \frac{\mathcal{T}\, T\, t_m}{N^2\,[\,1 - 2\,R(1 - \mathcal{T}/N^2) + (R^2 - T^2)\,(1 - \mathcal{T}/N^2)^2\,]}$$

Here $\mathcal{T}$ is the transmittance of the surface for uniformly diffused light incident from air, t_m is the transmittance of the surface for normally incident light, and R is the reflectance of the layer when immersed in a clear liquid of refractive index N. For a photographic plate, $N = 1.523$ for emulsion and glass, $\mathcal{T} = 0.905$, $t_m = 0.95$.

Both ϖ_0 and $\varkappa$ vary with wavelength; results of measurements at 546 mμ are given in Table 1 for 4 representative emulsions, together with approximate resolving powers for high contrast objects as generally quoted for similar plates.

The photographic effect of radiation is presumably dependent on the total intensity of radiation J, defined as $\int I\, d\Omega$, where the specific intensity I is the radiant flux per unit solid angle crossing unit projected area normal to the beam, and the integral is taken over a complete solid angle. The contribution dD to the density of the developed emulsion arising from a layer of thickness dz where the total intensity has been J for an exposure time t may be written

$$(2.2) \quad dD = f(J, t, z)\, dz.$$

Table 1

Properties of some photographic emulsions

Plate (Kodak, A'sia)	Lantern Plate Extra Contrast	Lantern Plate Normal	Process Panchromatic	Super XX
$\mathcal{R}$ (546 mμ)	0.79	0.72	0.41	0.28
ϖ_0 (546 mμ)	0.998	0.995	0.94	0.85
$T_{\frac{1}{2}}$	0 32	0.36	0.07	0.14
$\varkappa\zeta$ (approx.)	6	4.5	8	3
ζ (mm)	0.020	0.016	0.018	0.015
$\varkappa$ (mm^{-1})	300	280	450	200
	Orders of magnitude			
γ	5	2.5	5	1.2
γ_1 (mm^{-1})	2.5×10^2	1.5×10^2	2.8×10^2	8×10
Resolving power (lines/mm)	80	80	85	40

If the emulsion is fully developed throughout its depth the function $f(J, t, z)$ does not depend explicitly on z and, ignoring reciprocity failure, may be written $f(Jt)$. Further, over a wide range of illumination q, the density D of the developed emulsion is given closely by

$$(2.3) \quad D = \gamma \log\left(\frac{qt}{i}\right).$$

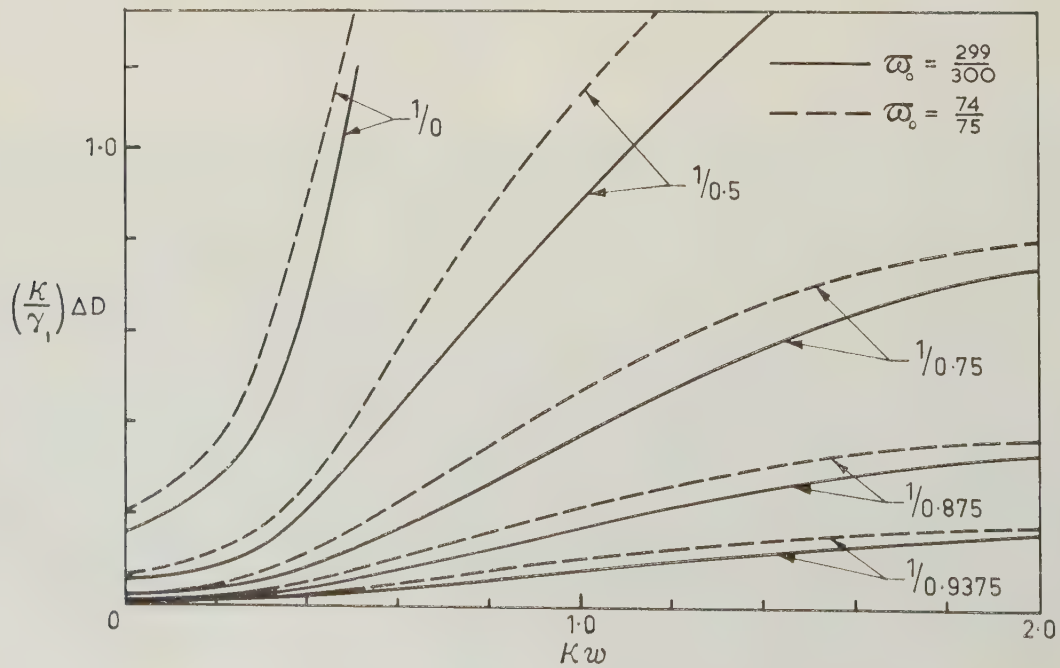

Fig. 1. — Variation of plate density ΔD with width w of a strip in the object pattern, and with object contrast. The full curves are for $\varpi_0 = 299/300$, the broken curves for $\varpi_0 = 74/75$.

where i is the plate inertia and γ a constant for the emulsion under standard conditions of development. The function $f(Jt)$ must be of a similar form

$$f(Jt) = \gamma_1 \log\left(\frac{Jt}{i_1}\right) \tag{2.4}$$

whence the density of the developed plate is at any point

$$D = \gamma_1 \int_0^{\zeta} \log\left(\frac{Jt}{i_1}\right) \mathrm{d}z. \tag{2.5}$$

To obtain the order of magnitude of γ_1, it will be noted that if J were constant with z,

$$D = \gamma_1 \zeta \log\left(\frac{Jt}{i_1}\right) \tag{2.6}$$

so that by comparison with (2.3)

$$\gamma_1 \approx \frac{\gamma}{\zeta}. \tag{2.7}$$

Typical values of γ_1 are also listed in Table 1.

III. **Diffuse radiation in photographic emulsions.** — Direct light falling on the emulsion is gradually attenuated by scattering and absorption, and the scattered light is re-scattered until it is either absorbed by grains or matrix or escapes from the emulsion. Thus there is a direct and a scattered component of the radiation in the emulsion, the latter being generally the more important.

Consider a co-ordinate system r, z whose r axis lies in the surface and is perpendicular to the length of the strips in the object pattern, z being the depth in the emulsion. If a long strip of surface at r, o, width $\mathrm{d}r$, is uniformly illuminated so that flux $2\pi q(r)$ enters the emulsion per unit area, the contribution to the total intensity at a point R, z due to diffuse radiation may be written

$$\mathrm{d}J_1 = q(r)\,\mathrm{d}r\,\Lambda_1(R - r, z).$$

If the whole surface be illuminated but with variations confined to the r direction, then

$$J_1 = \int_{-\infty}^{\infty} \Lambda_1(R - r, z)\, q(r)\,\mathrm{d}r. \tag{3.1}$$

The function $\Lambda_1(r, z)$ depends slightly on the direction of the incident light, but for most purposes it is sufficient to adopt the function for radiation entering the emulsion uniformly in a hemisphere. This has been given a formal expression by Giovanelli [3] for thick plane parallel diffusers. However, except near the bottom of the diffuser where values are relatively small anyhow, this function does not differ greatly from the simpler function $\Pi(r, z)$, which is the function $\Lambda_1(r, z)$ for a semi-infinite medium, and has been tabulated for a few cases [3].

In evaluating the contribution of direct radiation J_2 to the total intensity at depth z, it will usually be sufficient to assume collimated radiation incident normally on the emulsion, when the flux in the direct beam and J_2 are both given by

$$J_2 = 2\pi q(R)\, \mathrm{e}^{-\varkappa z}. \tag{3.2}$$

Adding both components, the total intensity is thus given closely by

$$J = J_1 + J_2 = \int_{-\infty}^{\infty} \Pi(R - r, z)\, q(r)\,\mathrm{d}r + 2\pi q(R)\, \mathrm{e}^{-\varkappa z}. \tag{3.3}$$

IV. **The resolving power of an emulsion.** — Some of the general features of the resolving power can be deduced immediately from the properties of the function $\Pi(r, z)$.

The total intensity in a semi-infinite diffuser differs from that in an infinite diffuser only because of the escape of light at the surface, near which the depletion of the radiation field is consequently greatest. Thus, except for small values of r, $\Pi(r, z)$ initially increases with depth for a given r, and falls off much more rapidly with r at the surface than at any given depth in the emulsion. This and the rapid decrease in direct light with depth are the explanation of the well known phenomenon that with fine patterns the resolved image resides near the surface of the emulsion [4].

A change in extinction coefficient alone, due for example to a change in the concentration of silver halide crystals, changes the resolving power by the same factor, for the physical and geometrical conditions are restored if the widths of the object pattern are changed in the same proportion.

The variation of $\Pi(r, z)$ with ϖ_0 or λ is rather intricate, but an increase in λ produced, say, by the addition of dyes, reduces the scale of scattered radiation, reduces its importance as compared with direct radiation, and therefore improves the resolving power.

As examples of the quantitative computation of resolving power, the variations in density ΔD across the plate have been computed from equations (2.5) and (3.3) for two fully developed emulsions, N(matrix) $= 1.5$, which have been exposed to a resolving power pattern. The albedos for single scattering are $\varpi_0 = 299/300$ and $74/75$ respectively, straddling the range of lantern plates. The integral in (2.5) has been taken only down to the depths $z = 0.75\,(\sqrt{3\lambda}\varkappa)^{-1}$ and $1.5\,(\sqrt{3\lambda}\varkappa)^{-1}$ respectively, beyond which the function $\Pi(r, z)$ has not been tabulated ; but at lower depths and for fine resolving power patterns the variation in $J_1(R - r, z)$ with R is fairly small and hence has little effect on the image contrast. The variations of $(\varkappa/\gamma_1)\,\Delta D$ with $\varkappa w$, where w is the width of a strip in the object pattern, are shown graphed for several values of the object contrast.

Results obtained for other types of emulsion

(i. e., for other values of ϖ_0) will show the same general features. For very narrow strips in the object pattern, the image contrast is due solely to the direct component of the total intensity, so that as ϖ_0 and the intensity of scattered light decrease the image contrast increases.

From the graph, the pattern reproduced with an image contrast corresponding to $(\varkappa/\gamma_1)\ \Delta D$ has a spacing given by the intersection of the line $(\varkappa/\gamma_1)\Delta D =$ constant with the appropriate curve for the given object contrast. There are two general cases : (1) The line intersects the curve for a black and white object pattern (contrast 1/0), giving the minimum pattern width which can be resolved with the required image contrast. (2) The line does not intersect the curve for a black and white object pattern, in which case patterns of all spacings are resolved for sufficient object contrasts. Since $\varkappa/\gamma_1$ lies in the range from about 2 to 5 for typical emulsions, ΔD must exceed 0.16/5 for extremely fine black and white object patterns for all normal plates. This corresponds to a variation in transmission of 7.5 per cent or more across the image pattern, which would be readily observable visually. However, in practice no plates resolve extremely fine patterns, and it would appear that density variations in the image are not the predominant limitation on resolving power for contrasty objects ; this must be attributed rather to granularity. For low contrast objects, the reverse must apply. If granularity sets a limit w_g to the finest pattern resolvable, and the limit for a given image contrast is w_c, the finest pattern actually resolved will be given closely by

$$w_r = \sqrt{w_g^2 + w_c^2}.$$

Arbitrarily selecting w_g and the image contrast at which the pattern is to be resolved, we can then obtain a family of curves against object contrast. With so many parameters available, ϖ_0, $\varkappa$ and w_g, as well as the image contrast for resolution, it is not very instructive to attempt to correlate the scanty observational results with theory, although with a suitable choice of parameters, curves can be predicted having the same general shape as those of Sandvik [5].

References

[1] R. G. Giovanelli, *Opt. Acta*, **2**, 1955, p. 153.
[2] E. Pitts, *Proc. Phys. Soc.* B, **57**, 1954, p. 105.
[3] R. G. Giovanelli, *Opt. Acta*, **3**, 1956, p. 49.
[4] T. H. James & G. C. Higgins, « Fundamentals of Photographic Theory », 1948, p. 257. John Wiley & Sons, Inc., New York.
[5] O. Sandvik, *Phot. Journal*, **68**, 1928, p. 313.

Manuscrit reçu le 15 août 1955.

A photon-counting, grazing incidence spectrometer for the 50-1 000 Å range

P. Fisher*, R. S. Crisp and S. E. Williams.
Department of Physics, University of Western Australia.

Summary. — *A vacuum spectrometer designed for investigating the soft X-ray spectra of metals is described. Gratings of one metre radius, 570 or 1150 grooves per mm are used at 85° angle of incidence. The detector is a Cu-Be photomultiplier, intensities being recorded in terms of pulse count rates on a potentiometer recorder. A method of preliminary calibration of the spectrometer, giving wavelength in terms of the position of the photomultiplier is described. Conditions for optimum resolution and definition of the spectra are discussed. Spectra of Mg, Al, Li, Be, Na, K and Cu are reproduced. Observations have been made of the effect of target contamination by oxidation and deposition of carbon. Cu M spectra are compared with those published by other workers and it is concluded that no sharp edges can be observed.*

Zusammenfassung. — *Es wird ein Vakuumspektrometer für die Untersuchung der weichen Röntgenstrahlung von Metallen beschrieben. Darin haben die Gitter einen Radius von 1 m, enthalten 570 oder 1 150 Striche pro mm und werden unter einem Einfallswinkel von 85° verwendet. Als Empfänger dient ein Cu-Be-Photomultiplier, wobei die Intensitäten als Impulszahlen mit einem Potentiometer-Registriergerät gemessen werden. Es wird ein Verfahren für eine vorläufige Eichung des Spektrometers beschrieben, die die Wellenlänge in Abhängigkeit von der Stellung des Photomultipliers angibt. Die Bedingungen für die beste Auflösung und Definition der Spektren werden diskutiert. Die Spektren von Mg, Al, Li, Be, Na, K und Cu werden wiedergegeben. Über die Wirkung einer Verunreinigung des Auffängers durch Oxydation und Ablagerung von Kohle wurden Beobachtungen angestellt. Die M-Spektren von Cu werden mit den Ergebnissen anderer Bearbeiter verglichen und die Folgerung gezogen, dass keine scharfen Kanten beobachtet werden können.*

Sommaire. — *On décrit un spectromètre conçu pour l'étude du spectre de rayons X mous des métaux. On utilise des réseaux de 1 m de rayon de courbure, de 570 ou 1 150 traits par mm et fonctionnant sous une incidence de 85°. Le détecteur est un photomultiplicateur au Cu-Be, les intensités étant enregistrées sous forme de fréquences d'impulsions sur un enregistreur potentiométrique. On décrit une méthode d'étalonnage préliminaire du spectromètre, donnant les longueurs d'ondes en fonction de la position du photomultiplicateur. On discute les conditions de résolution et de définition optima des spectres. On présente des spectres de Mg, Al, Li, Be, Na, K et Cu. On a observé l'effet de la contamination de la cible par oxydation et dépôt de carbone. Les spectres M du cuivre sont comparés avec ceux d'autres auteurs et on conclut qu'on ne peut pas observer des bords nets.*

A photomultiplier of the Allen type, using Cu-Be elements without a window, was first used as a detector in a grazing incidence spectrometer by Piore et al. [1]. In the investigation of the soft X-ray emission spectra of metals, photon-counting has the important advantage that the power dissipated in the target whose spectrum is excited by electron bombardment can be reduced almost a hundredfold in comparison with apparatus using photographic methods of recording. Since a spectrum can be recorded in a few minutes on a potentiometer chart compared with hours on a photographic plate, the photon-counting method allows the observation of spectrum changes produced, for example, by surface contamination of the target, which occur in time intervals of a minute or less. The effect of changing the position of the target or the focussing of the electron beam can be observed immediately. Finally, the range of intensities which can be measured between background and saturation with a multiplier is much greater than for a photographic plate.

The instrument described below was designed as a photon-counting spectrometer and has no provision for photography. An incidence angle of 85° is used with interchangeable gratings of 1 metre radius of curvature, one with 570 grooves per mm, the other with 1150 grooves per mm, ruled on aluminium-coated glass by the *National Physical Laboratory* and the *Nobel Institute* at Stockholm, respectively. The spectrometer is constructed almost entirely from austenitic stainless steel to improve the vacuum and magnetic properties. It has been used in the initial stages to record emission spectra from solid targets of a number of metals.

* Present address : Dept. of Physics, Purdue University, Lafayette, Indiana, U. S. A.

I. **Spectrometer.** — Figure 1 shows, diagramatically, the arrangement of the spectrometer, including the detector, the screw drive carrying the detector along the Rowland circle, and the source chamber. Power supplies, control circuits and pulse amplifying, counting and recording components are indicated by block diagrams. Figure 2 is a photograph of the interior of the spectrometer, showing the construction in more detail.

Radiation excited in the target by electron bombardment passes through the entrance slit, is diffracted by the grating and focussed on the Rowland circle. The detector, mounted on a carriage connected by a jointed arm to a nut on the lead screw is moved round the circle on rotation of the screw, its track being defined by two machined circular arcs. The analyser slit defines the beam of radiation falling on the photocathode of the multiplier. Pulses of electrons from the multiplier anode are led off by a busbar to a pre-amplifier, amplifier and counting rate meter which drives the pen of a recording potentiometer.

The two plates defining the circular groove, bearings for the lead screw, etc., are all bolted to an undertable which in turn is attached to the 16 1/2 inch diameter central plate. The diameter of the inner arc

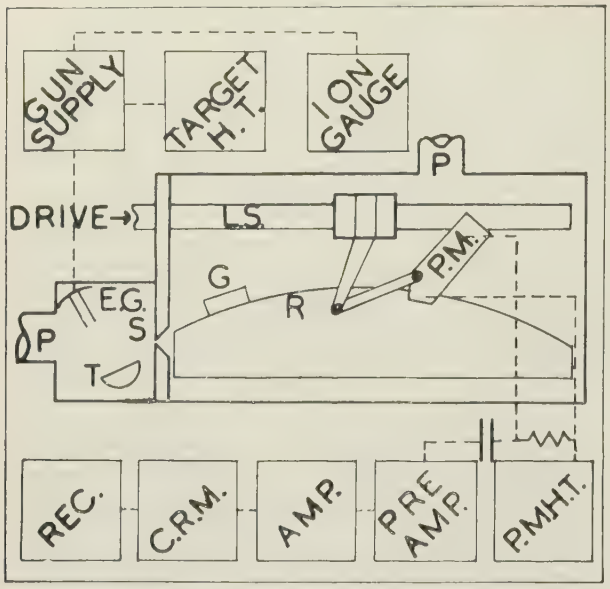

Fig. 1. — Arrangement of Spectrometer and associated electrical apparatus. R : Rowland circle ; G : Grating ; P. M. : Photomultiplier ; S : Source slit ; T : Target ; E. G. : Electron gun ; L. S. : Lead screw ; P. P. : Pumping ports ; C. R. M. : Counting Rate Meter.

is equal to the radius of curvature of the grating to within 0.001 inch and the plates are adjusted so that this arc, produced, passes through the source and analyser slits. The multiplier carriage runs on nine ball races to ensure accurate and smooth traversal of the Rowland circle. Two races are dowelled to the carriage and bear against the inner arc. Two more, sprung and pivoted, bear against the outer arc and keep the analyser slit on the Rowland circle. Four more are placed at the corners of the carriage and on these the assembly runs over the plates. It is held down firmly by the ninth roller which bears on the under side of the plate.

The carriage is driven by a $\frac{1}{4}$ h. p. motor, through a continuously variable hydraulic gear box, a 120 : 1 reduction and a bevel gear with a further train inside the vacuum. A bi-directional revolution counter and auxiliary dial allow estimation of the position of the lead screw nut to 0.001 turn or 0.0001 inch. A microswitch is tripped once in each revolution of the lead screw to put marker lines on the record and enable correlation of the observed spectrum with wavelengths calculated from the position of the analyser slit. All speeds up to 3 r. p. m. of the lead screw in either direction are possible.

The grating mount, housing a 2 inch diameter blank, must be as compact as possible. Three translational and three rotational motions are provided for adjustment, the rotation axes all passing through the pole of the grating. A second slit placed between source

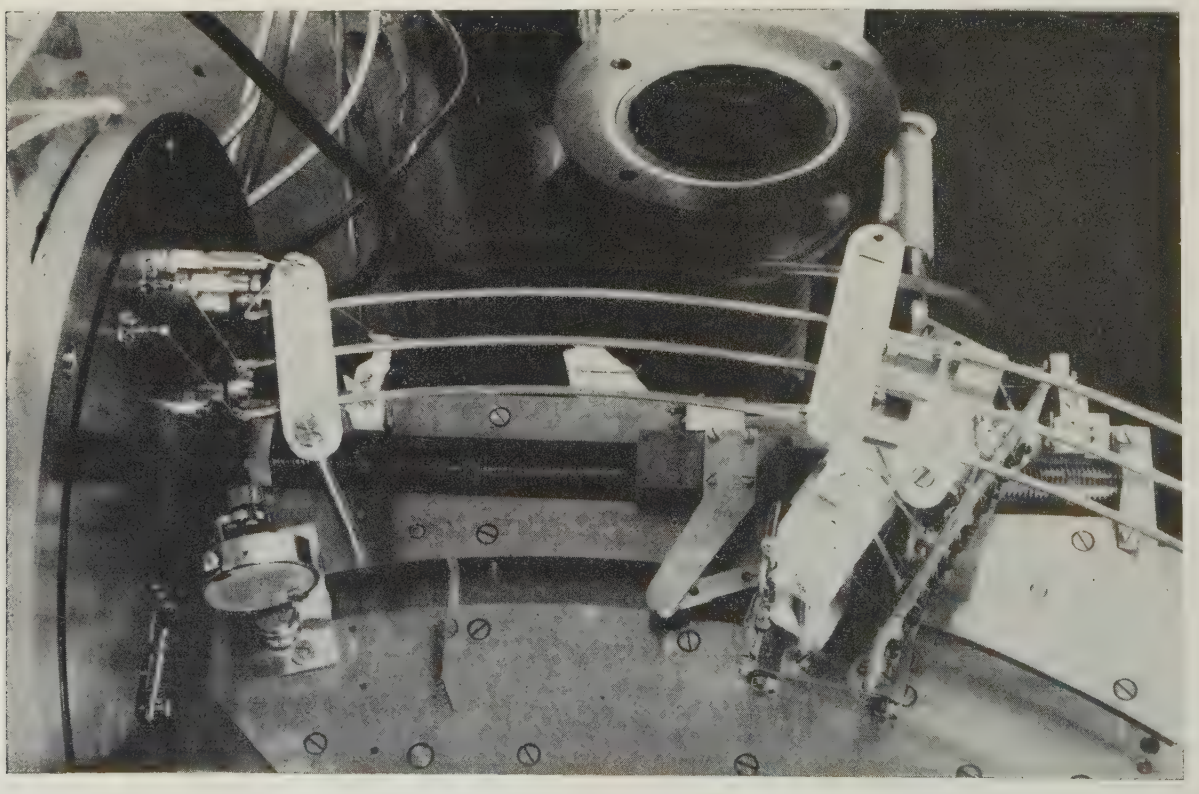

Fig. 2. — Interior of spectrometer.

slit and grating eliminates scattered radiation coming from the source chamber and limits the width of grating illuminated. Between the grating and the analyser slit there is a shield to eliminate the reflected ray.

The spectrometer chamber is evacuated by a 4 inch oil diffusion pump via a liquid air trap. Without cooling the trap a pressure of 10^{-5} mm, adequate for operation of the multiplier, is attained. Cooling with liquid air reduces the pressure to 10^{-6} mm which results in an appreciable decrease in target contamination, the contaminant vapour apparently entering the source chamber from the spectrometer through the source slit.

II. **Source chamber.** — The source chamber, figure 3, is in the form of a rectangular box 6 inches by 4 inches, with flanges welded to the open ends. One flange is bolted to the central plate, while to the other is attached an end plate carrying the pumping port, ionisation gauge, shaft for controlling the isolation valve and an evaporating furnace. All flange seals are made with " Silastic " O-rings.

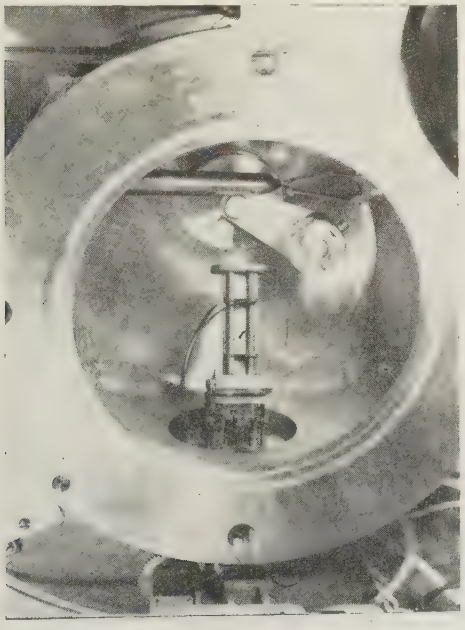

FIG. 3. — Interior of source chamber from front. Left, electron gun ; right, target ; below right, scraper. The isolation valve can be seen behind the target.

In the upper face of the source chamber is mounted the target holder, a hollow copper section, mounted on a system of O-rings and pyrex tubing. These allow adjustment of the position of the target without impairing the vacuum and also insulate the target from earth. The target is water cooled. Directly underneath is a rod operating through a double O-ring seal and carrying a tungsten carbide knife edge, spring mounted, which is used to scrape clean the face of the target. Specimens are retained in the target holder by dovetails and wedges. At the left side is the electron gun, which is similar to that described by KINGSTON [2], except that a tungsten spiral is used as filament. The gun unit is supported on a plate sealed to the wall by an O-ring.

The evaporating unit, a tantalum cup heated by electron bombardment is placed below the electron gun. It can be used without moving the target and, if necessary, evaporation can be continued while the target is being bombarded.

An isolation valve covering the source slit can be operated from the front of the chamber. Rotation of the shaft first lifts a plate from an O-ring seal surrounding the source slit, then turns the plate aside. This allows the source chamber to be opened when needed without letting air into the spectrometer.

The source chamber is evacuated by a 3-inch silicone oil diffusion pump via a re-entrant glass cold trap, seals being made by a system of O-rings and flanges. The source chamber working pressure is about 6×10^{-6} mm.

III. **Auxiliary electrical apparatus.** — The electron gun filament emission is stabilised and can provide a target current of 6 mA. The high tension for the target can be varied continuously from zero to 6 kV.

The high tension for the photomultiplier is derived from a stabilised r. f. circuit delivering 660 μ A at 5 kV. The voltage dividing system for the multiplier dynodes consists of a bank of 670 kΩ resistors, coated, with glyptal and baked, which are mounted on a pyrophyllite sheet on the multiplier block. These can be seen in figure 2. The resistors, which are functioning at 25 % of their rated power, have proved quite stable in vacuum. The h. t., signal and earth return are taken through seals in the central plate onto brass busbars with which phosphor bronze wipers mounted on the photomultiplier make satisfactory contact.

The pre-amplifier and amplifier are the model 500 [3] with a maximum gain of 5×10^5. Amplifier and multiplier h. t. connections, and the coupling to the input must be completely shielded. The counting rate meter is linear to better than 1 %. It incorporates a discriminator to reduce electrical noise and is modified to drive a " Sunvic " recording potentiometer.

Ionisation gauges in spectrometer and source chamber are combined with a safety relay which turns off gauge and electron gun filaments and diffusion pumps, should the pressure rise. Backing pressure is indicated by thermocouple gauges and the vacuum in the system is protected against failure of the backing pumps by magnetic safety valves [4].

IV. **Adjustment and calibration of the spectrometer.** — In figure 4, SS′, which is approximately parallel to the spectrometer axis, passes through the source slit and the focal point on the ROWLAND circle for 1000 Å. SS′ is actually a scribed line on the appropriate plate. The following adjustments must be made, jigs and a microscope being used for (i) and (ii), viz :

(i) The analyser slit axis is set on the machined Rowland circle ;

(ii) The circle and SS′ are adjusted to pass through the axis of the source slit ;

(iii) The grating is set on the circle at an arbitrary angle of incidence and adjusted while observing, with a microscope focussed on the analyser slit plane, the direct image of the source slit ;

(iv) The angle of incidence is set to 85° by placing the analyser slit at the calculated position for the direct image and moving the grating round the circle and rotating it about a vertical axis until the image viewed at the analyser slit is central and in best focus ;

(v) The second slit, set to a pre-determined width, is inserted to limit the grating aperture.

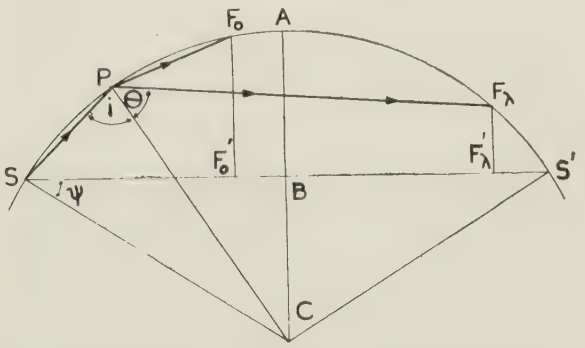

Fig. 4. — Geometrical diagram for preliminary calibration of spectrometer. S : source slit ; P : grating pole; F_0 : focal point for relected ray ; F_λ : focal point for wavelenght λ ; S′ : focal point for 1 000 Å.

Preliminary calibration involves the construction of a table of lead screw readings against wavelengths detected by the analyser slit. Figure 4 shows the geometry involved. SS′ has been defined. P is the grating pole and F_λ is the focal point for wavelength λ. If F_λ' is the foot of the normal from F_λ onto SS′, then SF_λ' is a_λ. From figure 4 it can be shown that $a_\lambda = R \sin (i + \theta) \sin (i + \theta - \psi)$, where i, θ and ψ are as indicated in the figure and R is the radius of curvature of the grating. From the grating equation, $n\lambda = \sigma (\sin i - \sin \theta)$, can be calculated values of θ for given λ and i. Using $i = 85°$ and ψ calculated from known dimensions, a table was compiled of a_λ against θ in 10 Å intervals from zero to 1000 Å.

With precision vernier callipers values of a_λ can be measured for each revolution of the lead screw. The two tables can then be combined to give the required calibration, i. e. lead screw reading against wavelength.

A correction curve for the calibration table was constructed by measuring the wavelengths of the emission edges of magnesium and aluminium in the first and higher orders using the table and the wavelengths quoted by Skinner [5] for these features. The observed wavelengths of the emission edges are consistent to within 0.05 Å. The correction varies from about 0.4 Å at 170 Å to 2.0 Å at 800 Å when using the 570 grooves per mm grating. The resulting dispersion curve is sufficiently precise for the purpose for which the instrument was designed, changes in the wavelength of, for example emission edges, being determinable to, at least, the above figure of 0.05 Å.

V. **Optical resolution and electronic definition.** — The spectrum produced is dependent on both optical and electronic performance of the instrument.

Since the analyser slit is normal to the Rowland circle, if s is the width of the source slit and s'' the width of the image of s focussed on the plane of the analyser slit, then $s'' = s/\sin \theta$, where θ is the angle of diffraction. Assuming that the width of the analyser slit, a, is less than s'', the response of the detector to truly monochromatic radiation is trapezium-shaped. Applying the Rayleigh criterion for resolution of two overlapping trapeziums of wavelength difference dλ, we have [6] $d\lambda = \left(s + \frac{a}{5} \sin \theta\right) \frac{\sigma}{Rn}$, where σ is the grating constant, R the grating radius, and n the order of the spectrum.

For optimum intensity s is made equal to a. For the 570 groove per mm grating, the width of slits is about 0.05 mm. For the 1150 groove per mm grating the width is about 0.08 mm. The resolution in the two cases, should then be 1.05 Å and 0.8 Å respectively. Although this does not match the resolution obtained by photographic techniques unless the third or fourth order can be used, a compromise must be made between resolution and intensity in order that counting statistics may be sufficiently good to ensure that usable and reproducible spectra are obtained.

When the spectrum is recorded as the detector is moved steadily round the Rowland circle, the speed of traverse must not exceed a value determined by the resolution and the response time of the recording apparatus. This response time includes the time which the counting rate meter takes to attain its final reading (within a small percentage) and the time which the recorder pen takes to move over the scale. The latter, for the recorder in use, is less than 2 seconds for the full scale and is negligible compared with the former except for very high counting rates of the order of 20,000 counts per second. Such high counting rates are quite unnecessary to good statistics and in practice are not attained.

The response times of the counting rate meter, for commonly used settings, are shown in table 1.

Table 1.

Counting rate range c. p. s.	Time in seconds for counting rate meter to attain final reading within :			
	3.5 %	2 %	1 ½ %	1 %
5 000	—	—	—	4
2 000	—	—	4	9
1 000	—	4.4	8	—
500	3.1	8.8	—	—
200	7,8	—	—	—

If the resolution is say, $d\lambda$, and the counting rate meter response time is t sec., then the speed of traverse should be such that $d\lambda$ is traversed in a time longer than t seconds. Depending on the counting range, the speed of traverse varies between about 5Å per minute and 12Å per minute. Figure 5 shows the effect of

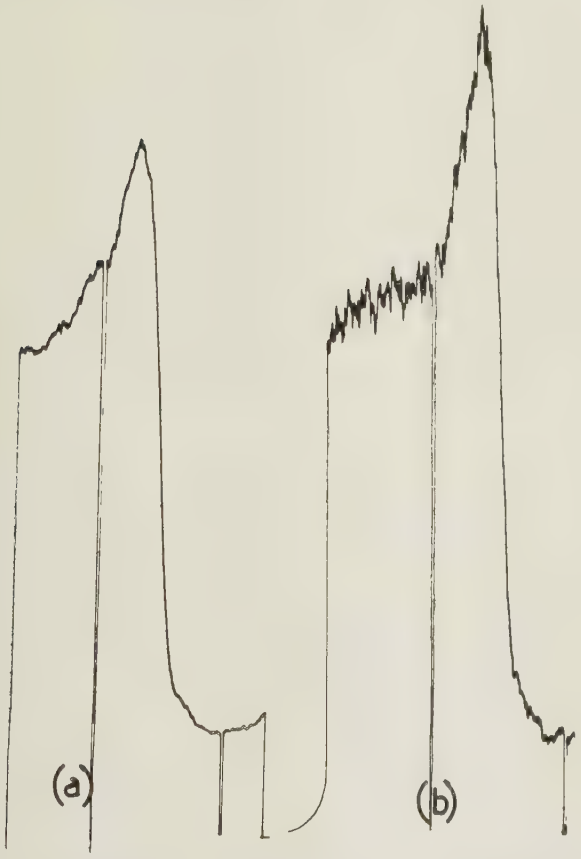

FIG. 5. — Traces of edge and peak of Mg $L_{2,3}$ spectrum. (a) Scanning speed too fast. (b) Correct scanning speed.

traversing the edge and peak of the magnesium $L_{2,3}$ band at a speed and with a response time (a) much greater than the optical resolution, and (b) with slower speed and shorter response time in accordance with the condition laid down above.

VI. **Performance.** — The effect of target contamination can easily be observed by recording the variation of intensity with time at a fixed wavelength. Figure 6 shows typical results for Mg, Al and Li. Following a rapid fall in the first 20-30 seconds after scraping (two minutes for Li) the decrease is nearly linear with time, the rate being independent of bombarding voltage. It is, however, dependent on the pressure in the source chamber, as the two Mg curves show. The difference in absolute intensity was due to different slit widths.

The initial rapid decrease in the case of Mg and Al is assumed to be due to the formation of an oxide layer. The linear decrease is due to the formation of a carbonaceous layer produced by the decomposition of residual vapours under electron bombardment, following their condensation on the target face. Contamination rates can be reduced almost to zero for Mg and Al by pumping for three or four days. Allowance for the effect of contamination on the shape of the spectrum can easily be made by observing twice at the

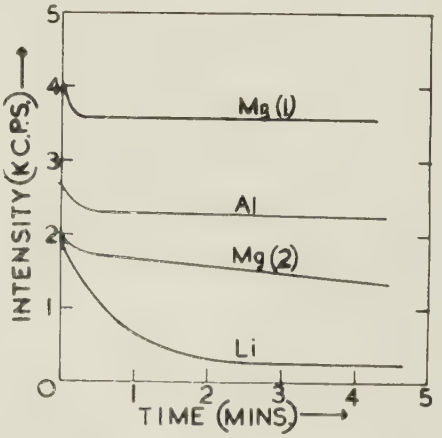

FIG. 6. — Variation of count rate with time due to target contamination. Mg(1) taken with pressure in source chamber 4×10^{-6} mm Hg ; Mg(2) at 8×10^{-6} mm Hg with a decreased slit width.

same wavelength, or more commonly, reversing the lead screw drive and recording the spectrum a second time in the opposite direction.

The K spectrum of carbon at 45 Å was looked for but was not seen, nor has the spectrum been seen in

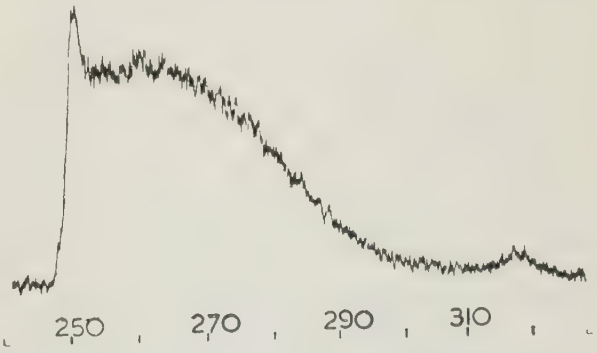

FIG. 7. — (a) Mg $L_{2,3}$, 2nd order, resolution 1 Å, Max. coun trate 1100 cps ; 2kV, 6 mA.

higher orders. It is assumed that the angle of incidence is low enough to eliminate it.

Magnesium has been used in spectroscopically pure form as a solid specimen to provide spectra from which standards of intensity and resolution could be obtained to judge the performance of the spectrometer at any time. Figure 7 shows (a) a second order L-spectrum taken with the 570 groove per mm grating, slit widths 0.1 mm, theoretical resolution 1 Å, and (b) a first order spectrum of the 1150 groove per mm grating,

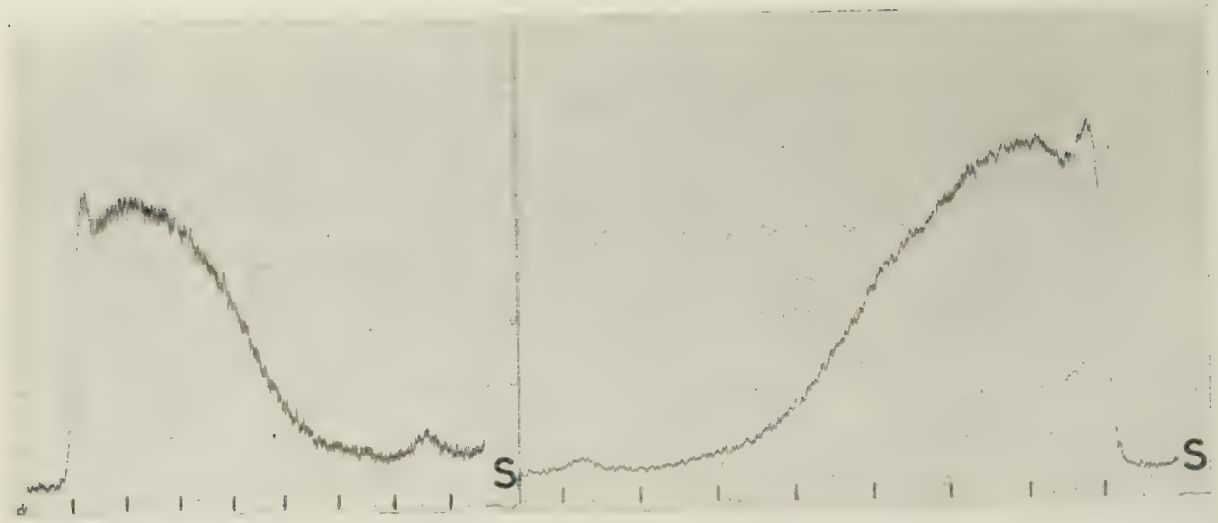

Fig. 7. — (c) Effect of contamination on shape of Mg spectrum. Record taken from right to left. Target cleaned by scraping at points marked S. Resolution 2 Å. Max. count rate 3 000 cps.

slit widths 0.08 mm and resolution 0.75 Å. The band width is about 7.0 eV. The ratio of height of peak to height of main band in (a) is 1.3 compared with 1.13 in the spectrum published by Skinner [5], the contamination rate in the present case being nearly zero.

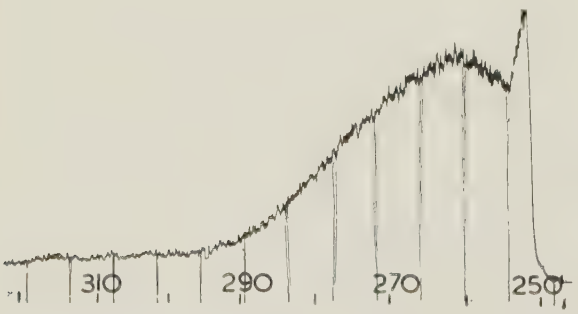

Fig. 7. — (b) Mg $L_{2,3}$, lst order, resolution 0.75 Å. Max. count rate 1 000 cps ; 4 kV, 6 mA.

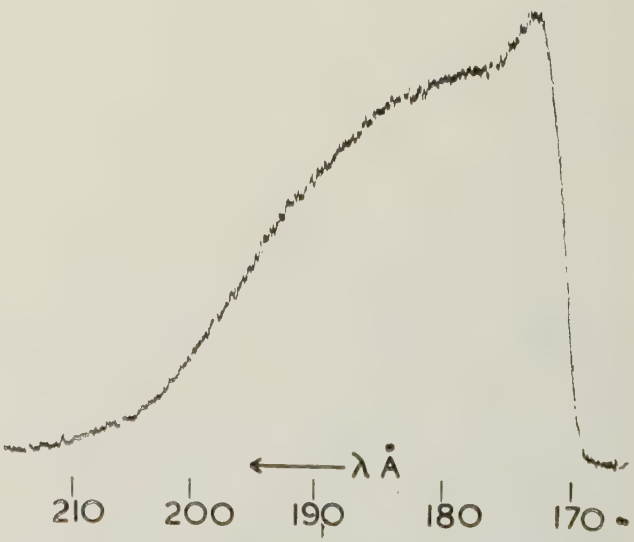

Fig. 8. — Aluminium, L spectrum. Resolution 2 Å. Max. count rate 7 000 cps.

Figure 7 (c) shows the effect of contamination on the shape of the spectrum, the specimen having been scraped and the drive reversed after recording the first spectrum. The scanning speeds are different for the two spectra. Figure 7 (c) also shows the Mg " line " at 318 Å after scraping the specimen surface. The effect of contamination is to reduce the intensity of the line in the same proportion as that of the band. It is concluded that the line is not associated directly with the presence of a contaminating layer. The Mg satellite band arising from double ionisation in the $L_{2,3}$ shell was observed with a bombarding voltage of 1850, the intensity being about 1.5 % of that of the main band.

Figure 8 shows a first order L-spectrum of aluminium taken with the 570 groove per mm grating and a theoretical resolution of 2 Å. The sharp spike at the high energy edge does not show and is not to be expected with this resolution. The band width is 13.3 eV.

The aluminium " line " has been observed at 275.3 Å compared with 290 Å observed by Skinner (1940) [5], the difference indicating that the energy level is to be attributed to an impurity rather than a surface state. Spectroscopically pure aluminium was used.

A first order K-spectrum of lithium, with resolution 1 Å, is shown in figure 9 (a). The shape confirms that observed by Skinner [5] and agrees with the measurements given by Bedo and Tomboulian [7]. The " roll-over " between edge and peak is indisputable. The solid specimens of lithium had been packed in kerosene to prevent oxidation. Under bombardment

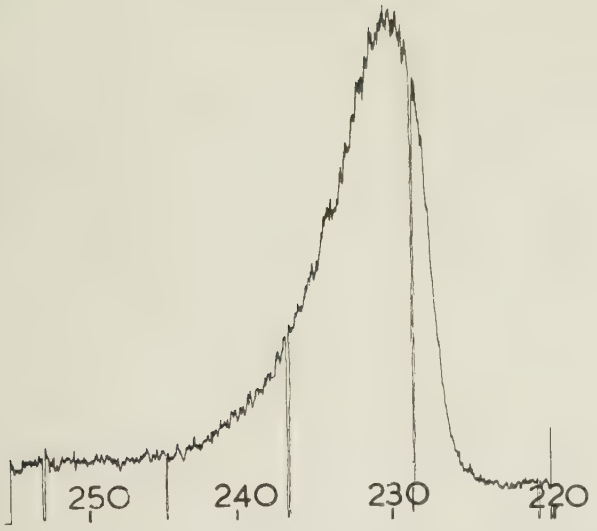

Fig. 9. — (a) Li K spectrum from vacuum-heated solid target. Resolution 1 Å. Max. count rate, 400 cps ; 4 kV, 6 mA.

the kerosene comes to the surface and a visible fluorescence is observed. The spectrum shown at (a) was from a vacuum-heated specimen and it shows no lithium " line ". In figure 9 (b) is shown a lithium spectrum with a very pronounced " line " and decreased intensity in the K spectrum, taken when kerosene was present on the surface. When lithium is evaporated onto the target, the " line " is, naturally, not obser-

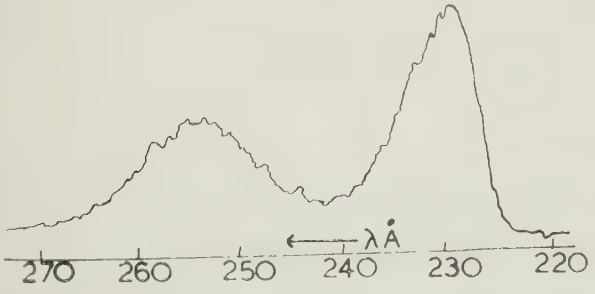

Fig. 9. — (b) Li K spectrum and Li " line " obtained with kerosene on target surface. Max. count rate 500 cps.

ved. Observation of the intensities of the " line " and the main band at various intervals after scraping has shown the " line " to be definitely due to surface contamination associated with the kerosene packing. The lithium band width is 4.2 eV.

A first order K-spectrum with resolution 1 Å, taken from a solid specimen of beryllium is shown in figure 10 (a). The band due to BeO at 122Å is not observed in a spectrum taken immediately after the specimen has been scraped. Figure 10 (b) shows that it becomes more prominent as contamination increases. The observed Be band-width is 12.7 eV.

The L-spectrum shown in figure 11 (a) was taken from a solid specimen of dry packed sodium. This

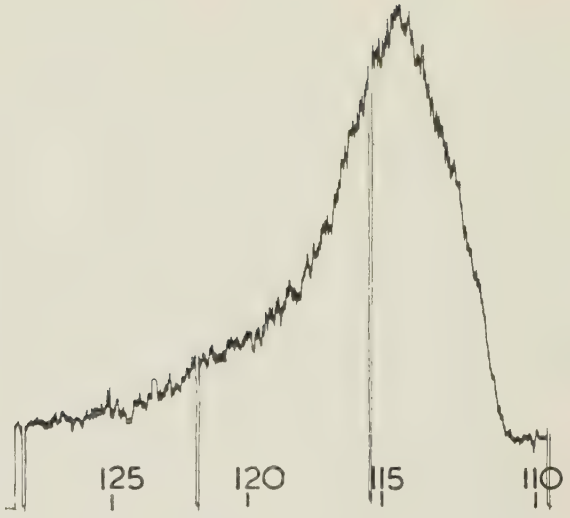

Fig. 10. — (a) Be K spectrum. Resolution 1 Å. Max. count rate 650 cps ; 4 kV, 6 mA.

was not scraped or cleaned except by electron bombardment, which was observed to remove the surface layer, leaving a clean metallic surface exposed. The

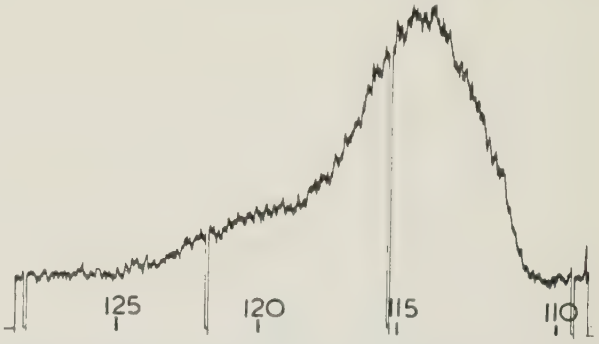

Fig. 10. — (b) Be K spectrum with BeO band due to target contamination by oxide. Resolution 1 Å. Max. count rate 450 cps ; 4 kV, 6 mA.

shape of the $L_{2,3}$ band, taking into account the resolution, 1 Å, agrees with the results from evaporated targets. The edge width is about 0.25 eV greater than can be accounted for by instrumental broadening.

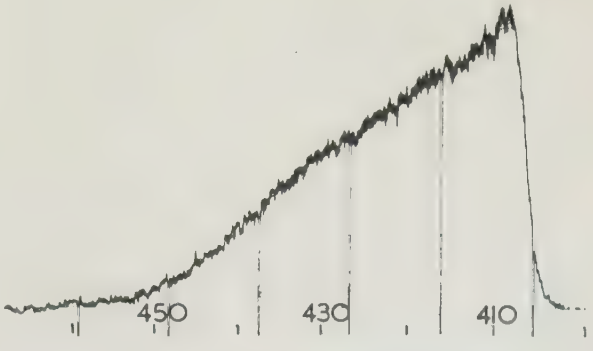

Fig. 11. — (a) Na L spectrum from solid target. Resolution 1 Å Max. count rate 1 800 cps ; 3 kV, 6 mA.

This may be due, at least in part, to temperature effects as in the case of soft alkali metals good thermal contact with the target holder was not attained and difficulty was experienced with targets which melted and even evaporated. The " line " at 375 Å recorded by Skinner has been observed. The band width is 3.3 eV.

In figure 11 (b) is a the M band of vacuum-melted potassium which had been spread onto the target with a knifeblade. No cleaning was possible, except by bombardment with the electron beam, which produced the same effect as with sodium. The spectrum agrees closely with that obtained by Kingston [2] from an evaporated layer. The band width is 2.2 eV.

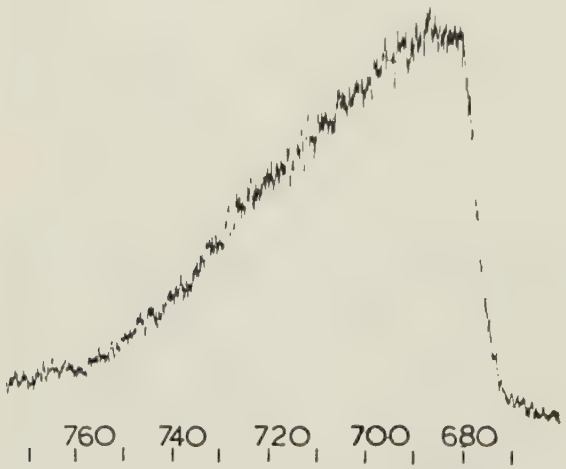

Fig. 11. — (b) Potassium M-spectrum. Resolution 1Å. Max. count rate 600 cps ; 4 kV, 6 mA.

Gyorgy and Harvey [8] published an M-spectrum of copper which they claimed showed M_2 and M_3 edges. Their result has been criticised by Skinner et al. [9] who found no edges in photographs of copper M-spectra and argued that the transition metals should not be expected to show definite edges. Figure 12 shows three copper curves. Curve (1) is enlarged from the published curve of Skinner, Bullen and Johnston [9]. The second curve (2) is the mean of seven separate curves obtained with the present instrument. Curve (3) is due to Gyorgy and Harvey, being replotted from their published curvilinear record. Seven curves were combined as the intensity is too low for good counting statistics and a single curve is not trustworthy as to shape. The mean curve (2) obviously agrees with curve (1) rather than curve (3). The band width is about 8.8 eV.

At the source chamber pressure of 4×10^{-6} mm, oxide formation was observed to reduce the intensity from a freshly scraped copper surface by 20% in four minutes after which time the intensity remained constant as very little further carbonaceous contaminant was formed.

Gyorgy and Harvey report that at bombarding voltages above 700 V the satellite becomes so intense that it obscures the emission edges. The seven spectra whose mean is plotted in (2) were taken with a bombarding voltage of 1850 V. The satellite is present with about the same intensity as in curve (3), but it is evident that it could not have obscured the emission edges if they were present. Interpretation of the shape of the Cu M-spectrum in terms of M_2 and M_3 bands, which is completely arbitrary since both separation and relative intensity of the components can be varied at will, has not been attempted.

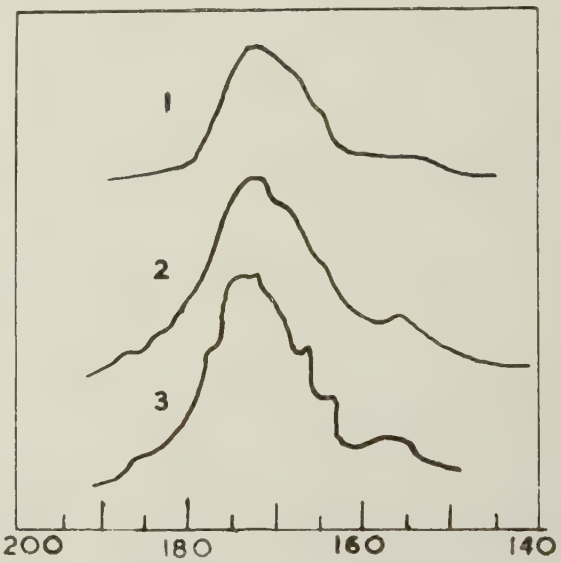

Fig. 12. — Cu M-spectra. (1) Reproduced from Skinner, Bullen and Johnston [9]. (2) Mean of seven spectra from present instrument. (3) Spectrum due to Gyorgy and Harvey [8], converted from original curvilinear to orthogonal co-ordinates. Max. count rate for (2) approximately 200 cps.

VII. **Conclusion.** — The results obtained justify the use of the photomultiplier detection technique, the performance of the instrument being consistent with calculation. They illustrate well the particular advantage of the technique in being able to record spectra in a relatively short time, to follow contamination processes and to observe the modifications which occur in the band shapes as these changes proceed. Future studies will be directed towards the band shapes and edge shifts in metallic solid solutions.

The authors wish to acknowledge the gift of two Cu-Be multiplier electrode systems by Dr J. S. Allen of the University of Illinois, the 570 grooves per mm diffraction grating by the Director of the *National Physical Laboratory*, stainless steel by the *Broken Hill Pty.* Ltd., and a grant for purchase of a potentiometric recorder by the *Electrical Research Board* of Australia. They are indebted to Professor C. J. B. Clews for his constant guidance and interest, and to Messrs D. W. Everson, R. Wright, J. R. Budge and J. Schreurs of the Physics Department technical staff for their skill and ingenuity. P. F. is indebted to

Imperial Chemical Industries of Australia and New Zealand for a Fellowship and to C. S. I. R. O., and R. S. C. to the University of Western Australia and the Australian Atomic Energy Commission for research grants, during the tenure of which the work was carried out.

References

[1] E. R. Piore, G. G. Harvey, E. M. Gyorgy and R. H. Kingston, *R. S. I.* **23**, 1952, 8.

[2] R. H. Kingston, Tech. Report n° 193, 1951, Res. Lab. of Electronics, M. I. T.

[3] W. C. Elmore and M. Sands, " *Electronics* ", McGraw-Hill, N. York (1949).

[4] P. Fisher, *R. S. I.* **27**, 1956, 5.

[5] H. W. B. Skinner, *Phil. Trans. Roy. Soc.* A**239**, 1940, 95.

[6] P. Fisher, *J. O. S. A.*, **44**, 1954, 665.

[7] Bedo and Tomboulian, *Bull. Am. Phys. Soc.* **1**, 1956, 334.

[8] E. M. Gyorgy and G. G. Harvey, *Phys. Rev.* **93**, 1952, 365.

[9] H. W. B. Skinner, T. G. Bullen and J. E. Johnston, *Phil. Mag.* **45**, 1954, 1070.

Manuscrit reçu le 18 avril 1957

Interferometric study of microscope objectives

Yukichi Ukita and Jumpei Tsujiuchi
Government Mechanical Laboratory, Tokyo, Japan.

Summary. — *The authors have constructed a specially designed microscope interferometer which can observe the image of a point source or periodic pattern simultaneously with wave aberration, and made a series of experiments to make clear the correspondence between the aberration and the definition of image. Spot diagrams and response functions are also calculated numerically from the wave aberrations. The importance of the method of illumination was ascertained in these experiments by comparing perfectly coherent and incoherent illuminations.*

Sommaire. — *Les auteurs ont perfectionné un interféromètre à microscope spécial permettant d'observer l'image d'une source ponctuelle ou d'un objet périodique en même temps que l'aberration d'onde ; ils ont fait une série d'expériences pour se rendre compte de la correspondance entre l'aberration et la définition des images. Les graphiques montrant les points d'intersection des rayons avec le plan image (the spot diagram en anglais), ainsi que les réponses en fréquence spatiale, ont été obtenus par un calcul numérique, basé sur l'observation de l'aberration d'onde. L'importance du mode d'éclairage a été démontrée au cours de ces expériences en comparant des éclairages parfaitement cohérents et incohérents.*

Zusammenfassung. — *Mit einem hierfür besonders ausgeführten Mikrointerferometer, das die gleichzeitige Beobachtung der Bilder eines Sternes oder einer regelmässigen Strichgruppe und der Wellenaberration erlaubt, werden Versuche gemacht, um den Zusammenhang zwischen der Bildgüte und den Aberrationen zu klären. Die Zerstreuungsfiguren und die Kontrastübertragungsfunktion werden auch aus den Wellenaberrationen berechnet. Aus den Versuchen ergibt sich auch die Bedeutung der Beleuchtungsart durch den Vergleich rein kohärenter und inkohärenter Beleuchtung.*

1. Introduction. — There are two points of view from which the evaluation of an optical image can be carried out. For the first, in which the object is considered as the summation of the point sources, the image of a point source would be sufficient to characterize the performance of the optical system. For the second, the object is considered as the summation of a set of sinusoidal components, and the image of this component pattern would be useful for the purpose [1].

Another representation of the performance of the optical system is its aberration, and particularly in the optical designing the aberration is the only one. So it is very useful to clarify the correspondence between the aberration and the image of a point source or periodic pattern.

For this purpose, the microscope interferometer of the type of Twyman [2] and Bracey [3] is constructed, suitably modified for observing the image of a pinhole or Siemens star simultaneously with the wave aberration.

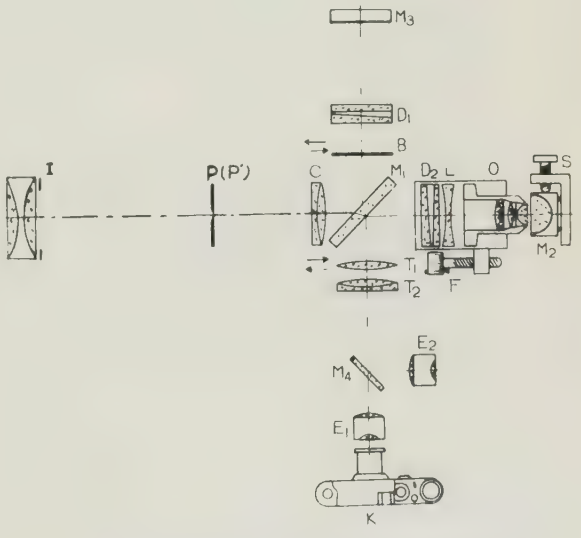

Fig. 1. — The microscope interferometer.

2. The microscope interferometer. — The schematic diagram of the microscope interferometer is shown in figure 1. P is a pinhole at the focus of a collimator lens C, and it can be replaced by a Siemens star P'. M_1 is a beam splitter and D_1, D_2 are deflectors. The objective to be tested is placed at O with a tube length lens L which makes the image point finite, and both can be moved by a fine screw F along the optical axis. Beams split by M_1 are reflected by M_3 and M_2 respectively. T_1, T_2 and E_1 (for photographic observation by a camera K) or E_2 (for visual observation) form a low power microscope which is focused at the exit

pupil of O. When T_1 is removed, the system becomes telescopic and the image of P or P′ formed by O can be observed, and the obstructive image formed by M_3 is stopped by closing a shutter B in this case.

Observations are made by monochromatic light of 546 mμ from a high pressure mercury lamp with Wratten 77 A filter, and P or P′ is illuminated through the condenser lens with iris stop I. By varying the size of the diameter of I, coherence of the illumination can be varied.

Deflection angle δ and azimuthal angle Φ of deflection of the reference beam can be varied by D_1. If δ is given suitably, the interference fringe itself can be taken as a diagram of wave aberration [4]. By varying Φ, the aberration of meridional, sagittal or any skew section is measured.

The testing beam is inclined to the axis of the objective to be tested by D_2, and M_2 can be moved perpendicularly to the axis by S. M_2 is a spherical mirror of radius 10 mm and of height 10 mm plus 0.17 mm, 0.17 mm is the thickness of the standard cover glass, as it is necessary to take the thickness of cover glass into consideration when dry objectives are tested.

3. Wave aberration and star image. — In figure 2, L is a wave front just emerged from the exit pupil of the optical system, and t is principal ray. Let us consider a point R(r, Φ) on the exit pupil with the polar

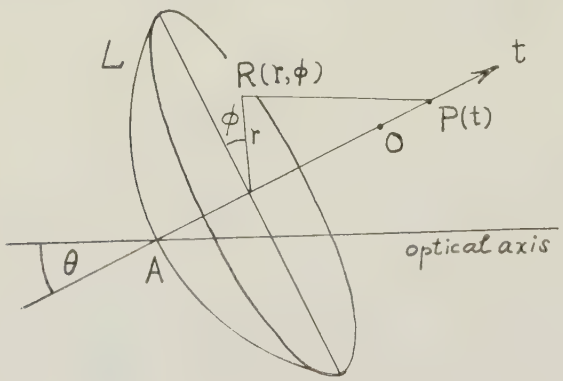

Fig. 2. — Wave front emerged from optical system.

coordinate (r, Φ), where Φ is measured from the intersection of the exit pupil and meridional planes. Let O be the Gaussian image point and A the intersection of the principal ray with the optical axis ; the wave aberration is the optical path difference between L and the reference sphere with O as centre and $\overline{AO}$ as radius.

Taking the reference sphere of radius $\overline{AP}$ centered at the focused point P(t), which is on the principal ray and a distance t from O, the wave aberration at P(t) is as follows :

$$(1) \qquad W(r, \Phi : t) = W_d(r : t) + W_0(r) + W_1(r)\cos\Phi + W_2(r)\cos^2\Phi + W_3(r)\cos^3\Phi + W_4(r)\cos^4\Phi + \ldots,$$

where W_d is the aberration due to the focal shift of P(t) from O,

$$(2) \quad W_d(r : t) = (1 - \cos\alpha)\, t, \quad \alpha = \sin^{-1}(r/\overline{AP}).$$

W_0 is independent of Φ and called the spherical type aberration ; $W_1(r)\cos\Phi$, $W_3(r)\cos^3\Phi$, ... are 1 st, 2 nd, ... order comatic type aberrations, and $W_2(r)\cos^2\Phi$, $W_4(r)\cos^4\Phi$, ... are 1 st, 2 nd, ... order astigmatic type aberrations, respectively [5].

If P(t) and Φ are given, the corresponding wave aberration curve is obtained from the interferogram. When $\Phi = 0^o$, the point of minimum residual aberration gives the meridional best focus P(t_m). The sagittal best focus P(t_s) and the best focus of any skew section can also be found in the same way. A point where the mean value of the residual aberration becomes minimum, is defined here as the best focus P(t_o), which corresponds to the position of the circle of least confusion.

The star images of two objectives of 40 × (N. A. = 0.65) have been studied at the successive focused points by comparing their wave aberrations. As the Gaussian image point is difficult to determine from the interferograms, the origin of t is taken at the best focus t_o for the axial image, and at the sagittal best focus t_s in case of an extra-axial image.

In figure 3, the star images and the corresponding interferograms of an objective are taken near the best focus of the axial image ($\theta = 0^o$). The interference fringe through the centre of the aperture, c in the figure, is the wave aberration curve itself. The distance between successive fringes corresponds to 1λ aberration. When $t' = t - t_0$ becomes as large as 0.003 mm, although the image of a pinhole is still a small point, a strong halo appears and contrast becomes low. The wave aberration in this case is zero about an axial zone but increase rapidly at the marginal zone. When $t' = 0$, the image is more or less dim, but its contrast is very high. At this position the residual aberration is minimum, and this is the position of best focus.

Figure 4 shows the images of another objective when $\theta = 4^o$. Meridional and sagittal aberrations are shown with interferograms, and from these figures the best focus ($t' = t - t_s = -0.002$ mm), the meridional best focus ($t' = -0.003$ mm) and the sagittal best focus ($t' = 0$) are determined.

In (1), neglecting the terms higher than $\cos^5\Phi$ and replacing r by the numerical aperture z and taking P(t_s) as origin, W becomes as follows :

$$(3) \quad W(z.\ \Phi : t) = W_d(z : t - t_s) + W_0(z : t_s) + W_1(z)\cos\Phi + W_2(z)\cos^2\Phi + W_3(z)\cos^3\Phi + W_4(z)\cos^4\Phi.$$

By using the three interferograms of the later objective taken at $\Phi = (0^o, 180^o)$, $(45^o, 225^o)$, $(90^o, 270^o)$ at three positions, t_s, t_o, t_m, when $\theta = 4^o$, the coefficients W_d, W_0, W_1, W_2, W_3, W_4 in (3) were obtained.

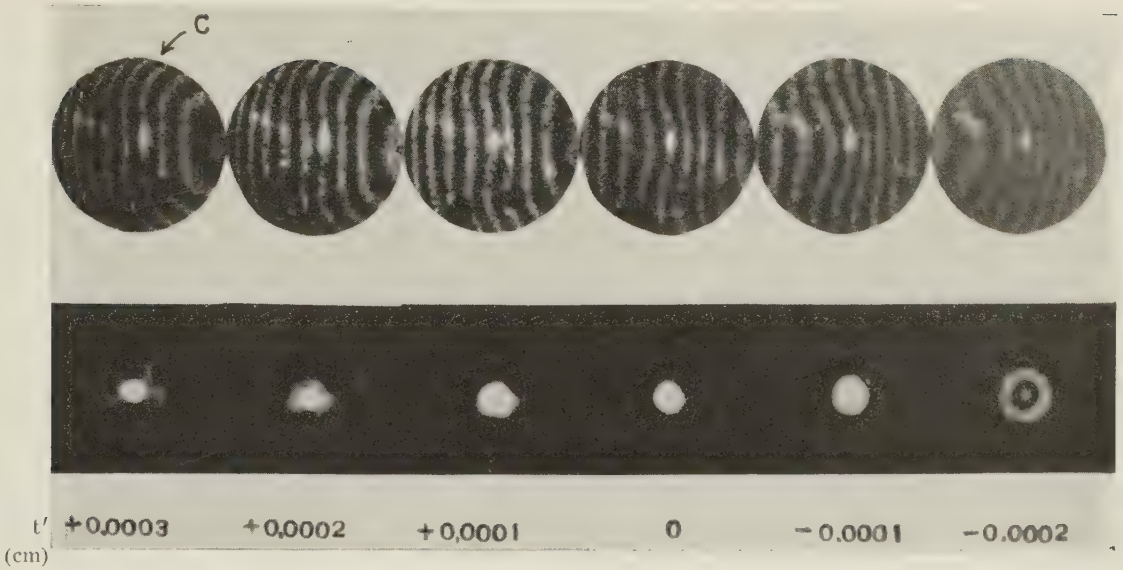

FIG. 3. — Star images and interferograms of a 40 × (N. A. = 0.65) objective at the successive focused points on the optical axis (θ = 0°).

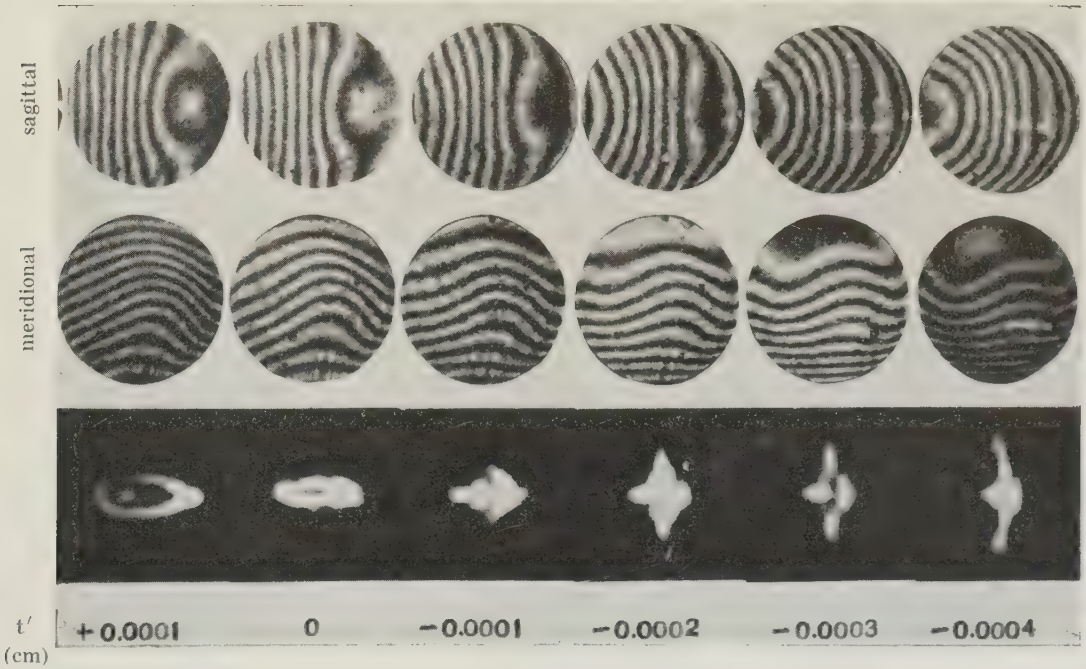

FIG. 4. — Star images and interferograms of a 40 × (N. A. = 0.65) objective at the successive focused points on the principal ray of θ = 4°

These results are shown in figure 5. Thus, the wave aberration of this objective when θ = 4° is perfectly determined.

Now let us try to show the correspondence between aberration and star image by spot diagrams [6]. A ray, passing through R(r, Φ), intersects the image plane which is perpendicular to the principal ray at P, and the intersection point P′ is specified by the (u, v) coordinate which has the origin at P and taking the intersection with the meridional plane as v-axis.

Let f be radius of reference sphere and N the refractive index of the image space, the numerical aperture z of the system becomes Nr/f, then the coordinates of P′ are [7]

$$(4) \qquad \begin{cases} u = -\sin\Phi \dfrac{\partial W}{\partial z} - \dfrac{\cos\Phi}{z} \dfrac{\partial W}{\partial \Phi}, \\ v = -\cos\Phi \dfrac{\partial W}{\partial z} + \dfrac{\sin\Phi}{z} \dfrac{\partial W}{\partial \Phi}. \end{cases}$$

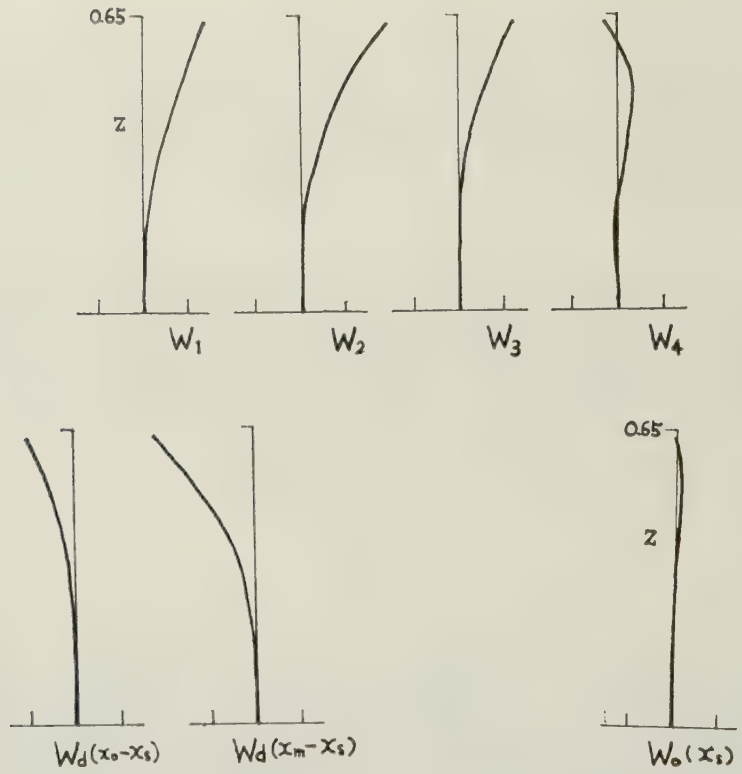

Fig. 5. — Wave aberration coefficients of a 40 × (N. A. = 0.65) objective when $\theta = 4°$ (abscissa unit ; 1λ).

Dividing the pupil into small elementary areas, $\frac{\partial w}{\partial z}$ and $\frac{\partial w}{\partial \Phi}$ are obtained from figure 5 and (3), so $P'(u, v)$ can be calculated from (4). By dividing the pupil into 2604 equal elements, we have calculated the spot diagrams numerically. The results are shown

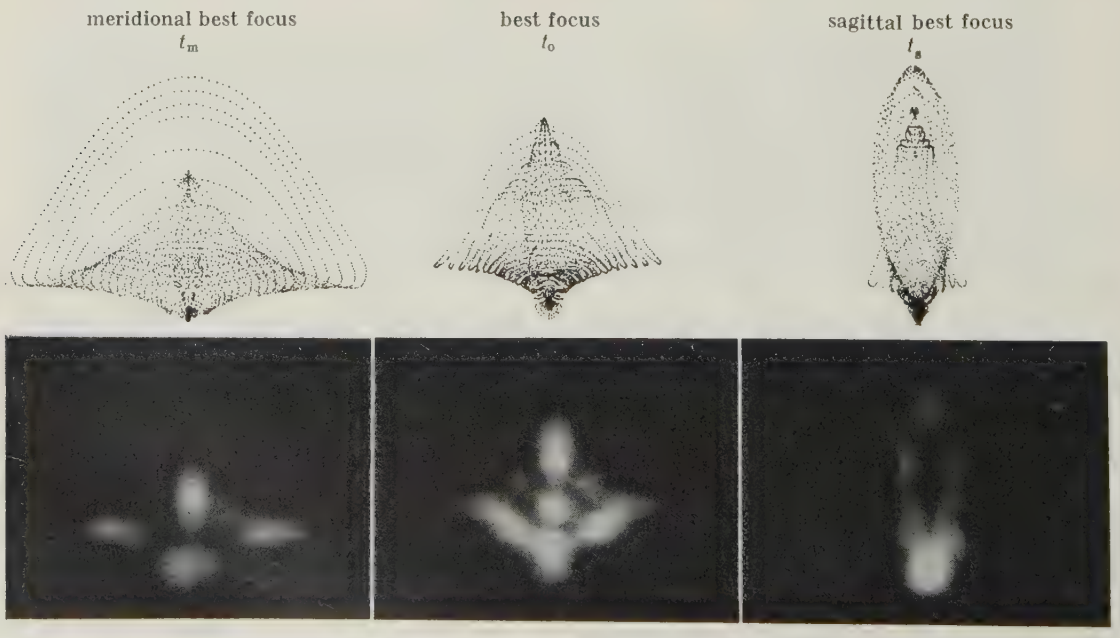

Fig. 6. — Spot diagrams and corresponding star images when $\theta = 4°$.

in figure 6 with the photographs of the star images for reference. Intensity distributions of the star images agree reasonably well with the spot diagrams.

4. Wave aberration and Siemens star image. — In order to study the image-forming characteristics of an optical system, it is most reasonable to use the response function (transmission factor) [8] [9], and so called resolving power can be estimated from it. For microscope objectives, the coherence of the illumination influences the results remarkably. The cases of the central coherent illumination and the incoherent illumination were studied.

1. Central coherent illumination. — If the numerical aperture of the condenser coaxial with the objective is very small as compared with that of the objective, it is called central coherent illumination.

According to Hopkins [9], if the complex amplitude of the test pattern is written as

$$E(u) = \frac{1}{2} + \frac{1}{2}\cos\ \omega u\ , \tag{5}$$

then the intensity of it is

$$I(u) = \frac{3}{8} + \frac{1}{2}\cos\omega u + \frac{1}{8}\cos 2\omega u\ , \tag{6}$$

and that of the image is

$$I'(u') = \frac{3}{8} + C(\omega,0)\frac{1}{2}\cos\omega u' + \frac{1}{8}\cos 2\omega u', \tag{7}$$

where u and u' are the normalized coordinates of the object and the image plane respectively and ω is the spatial frequency.

$C(\omega, 0)$ is the response function of fundamental frequency. For the axial image, $W(x, y)$, the wave aberration in terms of rectangular coordinates in the pupil plane becomes axial symmetric and $C(\omega, 0)$ is real. If W is normalized with respect to (x,y) in the frequency plane, then

$$C(\omega, 0) = \cos\,[\,k\,W(r)\,]\ , \tag{8}$$

where $k = 2\,\pi/\lambda$, λ is the wavelength, and

$$r = (x^2 + y^2)^{\frac{1}{2}}.$$

In our interferometer, the central coherent illumination is realized by making the iris stop I (fig. 1) very small. As $W(x,y)$ is measured by the interferogram, $C(\omega, 0)$ can be calculated by (8).

Let R be the number of lines per unit length in the test pattern, the spatial frequency ω corresponds to $(\lambda/z)R$. If we set a Siemens star P′ in place of P in the interferometer, the images of various R values can be observed at the same time.

Figure 7 and figure 8 show an example of $C(\omega, 0)$ and a Siemens star image formed by the objective of 40 × (N. A. = 0.65). Figure 7 (a) is the wave aberration obtained from the interferogram, figure 7 (b) the computed $C(\omega, 0)$ and figure 8 (b) is its image, respectively. For comparison the ideal image obtained by the reference beam (ray reflected by M_3) is shown in figure 8 (a).

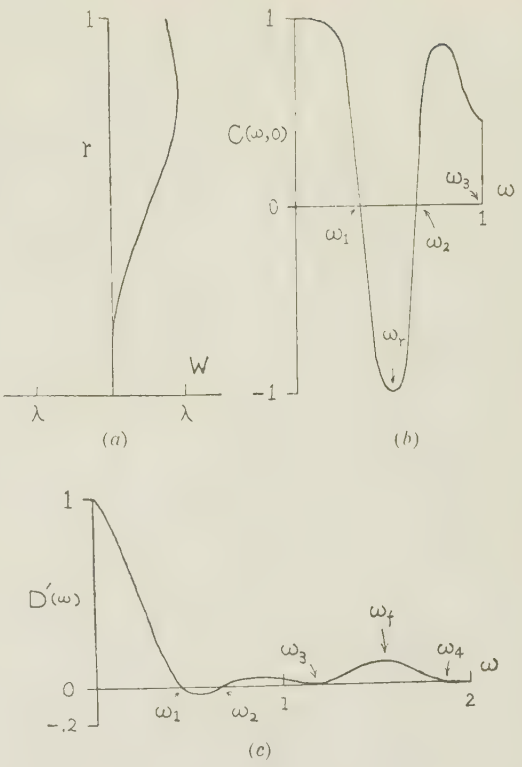

Fig. 7. — Response functions of a 40 ×(N. A. = 0.65) objective.

(*a*) Wave aberration; (*b*) $C(\omega,0)$, the response function of central coherent illumination ; (*c*) $D'(\omega)$, the response function of square topped pattern of incoherent illumination.

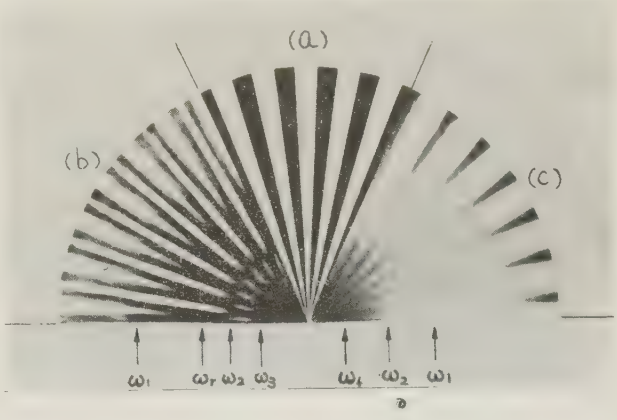

Fig. 8. — Siemens star images of the objective of figure 7.

(*a*) Ideal image ; (*b*) Image of central coherent illumination ; (*c*) Image of incoherent illumination.

It is obvious that the fundamental frequency, the second term in (7), makes so great contribution to the image, that the image varies greatly with the change of its response function. Near the first zero of $C(\omega, 0)$,

where $\omega_1 = 0.35$ ($R_1 = 400$ lines/mm), the second term of (7) becomes noticeable and it appears as if the pattern were that of the twofold frequency. At the region where $C(\omega, 0)$ is negative, the contrast of the pattern reverses, and $\omega_r = 0.51$ ($R_r = 600$), where $C(\omega, 0)$ is minimum, a clearly reversed pattern is observed. At the second zero of $C(\omega, 0)$, where $\omega_2 = 0.62$ ($R_2 = 740$), the pattern of twofold frequency is also observed. Beyond this frequency $C(\omega, 0)$ becomes positive again and the contrast of the pattern returns to the former states. At $\omega_3 = 1$ ($R_3 = 1200$), the contrast of the pattern vanishes suddenly, and beyond this frequency it can no longer be resolved just as the theory shows.

Experiments were carried out with three objectives

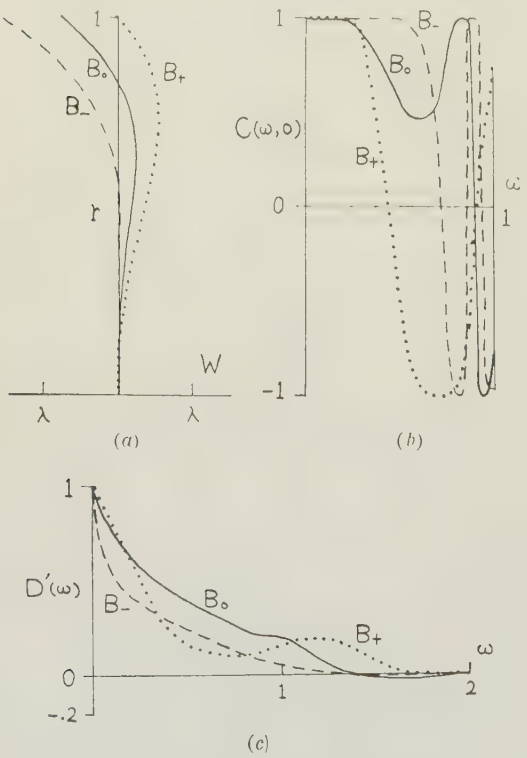

Fig. 9. — Response functions of a 40 × (N. A. = 0.65) objective " B " at the successive focused points.

(*a*) Wave aberrations ; (*b*) $C(\omega,0)$, the response functions of central coherent illumination ; (*c*) $D'(\omega)$, the response functions of square topped pattern of incoherent illumination.

of 40 × (N. A. = 0.65) ; they are denoted as " A ", " B ", " C ". The aberration of " A " is very small, " B " is moderate and " C " is very large. In figure 9 and figure 10, the response functions and the Siemens star images of " B " are shown at three points near the best focus. The sign + (—) means that the objective is moved by 0.001 mm *backward* (*forward*) towards the image plane from the originally focused point B_0. Figure 9 (a) is the wave aberration, figure 9 (b) the computed $C(\omega, 0)$, figure 10 (b) the Siemens star image and figure 10 (a) the ideal image.

The frequencies where $C(\omega, 0)$ becomes zero calculated and observed from Siemens star images in terms of R lines/mm of these three objectives are shown in Table 1.

Table 1. — Resolution limits of microscope objectives of 40 × (N. A. = 0.65) with 546 mμ in lines/mm.

		Central coherent		Incoherent	
		calculated	observed	calculated	observed
A_-	R_1	750	750	840	900
	R_2	1 090	960	2 000	—
A_0	R_1	1 200	1 200	2 400	2 200
A_+	R_1	600	480	550	600*
	R_2	1 100	800	1 550	1 350
	R_3	1 200	1 200	2 300	1 850
B_-	R_1	810	650	1 400	1 200
	R_2	960	720	—	—
	R_r a)	670	450	—	—
B_0	R_1	1 000	800	1 500	1 150
B_+	R_1	500	410	1 800	1 400
	R_2	1 000	780	—	—
C_-	R_1	700	570	1 700	800
C_0	R_1	390	320	1 450	1 100
	R_2	720	600	—	—
C_{++}	R_1	310	160*	340	—*
	R_2	560	420	1 050	930
	R_3	800	610	1 250	1 100

— absent, * doubtful, a) first minimum of response function.

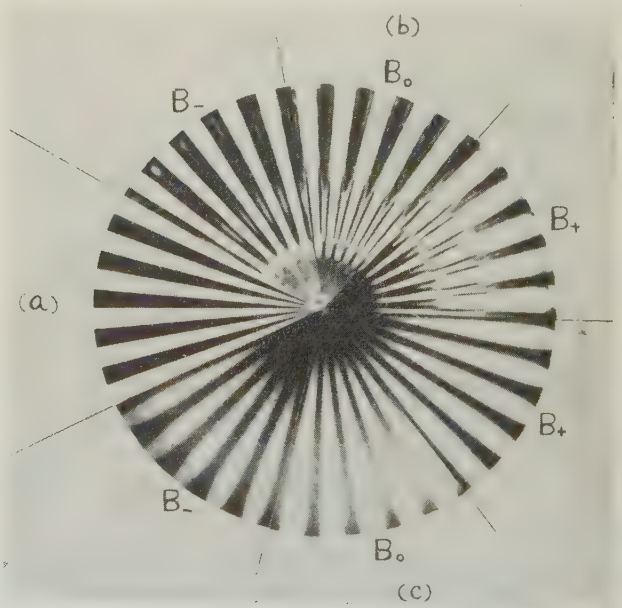

Fig. 10. — Siemens star images of the objective " B "

(*a*) Ideal image ; (*b*) Images of central coherent illumination ; (*c*) Images of incoherent illumination.

From these results, although it is seen that the pattern whose $\omega \leqslant 1$ are *resolved*, the contrast varies extremely and even reversal may occur when the amount of wave aberration is large. As reversed contrast is very harmful for usual purposes, it is reasonable

for the first zero of $C(\omega, 0)$ to be defined as the resolution limit of the objective. They are shown as R_1 in Table 1. Thus the SIEMENS star may be used to determine the resolving power of microscope objectives. It is clear from the above description that the monofrequency grating is unsuitable for the determination of the performance of microscope objectives.

2. Incoherent illumination. — When the numerical aperture of the condenser is very large as compared with that of the objective, or the object is self-luminous, illumination is incoherent. In this case, the objective is considered as a linear filter as to the intensity, and the response function becomes ;

$$(9) \qquad D(\omega) = \frac{1}{\pi} \iint_{-\infty}^{\infty} \exp \mathbf{i} \left\{ k \left[W\left(x + \frac{\omega}{2}, y\right) - W\left(x - \frac{\omega}{2}, y\right) \right] \right\} dx\, dy.$$

If the object be the same as that of the coherent illumination, i.e. (6), the intensity of the image is

$$(10) \quad I'(u') = \frac{3}{8} + D(\omega)\frac{1}{2}\cos \omega u' + D(2\omega)\frac{1}{8}\cos 2\omega u'.$$

If the object is a perfect square topped pattern, (9) is replaced by [10]

$$(11) \quad D'(\omega) = \frac{4}{\pi} \sum_p \frac{1}{p} D(p\omega), \quad p = 1, \quad 3, 5, \quad -7, \ldots.$$

The analytical calculation of $D(\omega)$ from the wave aberration was made by HOPKINS in the case of defocus [11], in the case of astigmatism by DE [12] and in the case of very small aberration by STEEL [13].

When the aberration is symmetrical about the axis, $D(\omega)$ becomes real, and then (9) becomes

$$(12) \qquad D(\omega) \quad \frac{1}{\pi} \iint_A \cos \left\{ k \left[W\left(x + \frac{\omega}{2}, y\right) - W\left(x - \frac{\omega}{2}, y\right) \right] \right\} dx\, dy,$$

where A is the integrating region common to the unit circles representing the aperture in the frequency plane. The numerical integration of this equation is carried out as follows :

In figure 11, let A_{ab} be elementary area in the integrating region distant by a and b from $E\left(-\frac{\omega}{2}, 0\right)$ and $E'\left(\frac{\omega}{2}, 0\right)$ respectively, the centres of unit circles representing the aperture. If we put

$$(13) \qquad B_{ab} = A_{ab} \cos \left\{ k [W(a) - W(b)] \right\},$$

(12) becomes

$$(14) \qquad D(\omega) = \frac{1}{\pi} \sum_a \sum_b B_{ab}$$

We divided the pupil into ten concentric zones of equal width of 0.1 by the circles centered at E and E', and calculated A_{ab} numerically from the given a, b and ω [14].

If the amount of variation of W in the zone cannot be neglected, the following treatment is preferable. Taking a (ξ, η) coordinate with $O(a, b)$, the centre of the elementary area, as origin, ξ, η axes are taken along the tangents to these two circles limiting the elementary area (fig. 11). Then W can be represented within the elementary area by

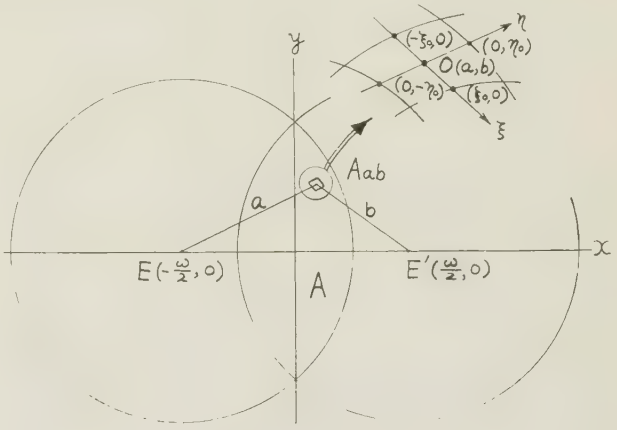

FIG. 11. — Elementary areas in the inte-grating region.

$$W(\xi) = W(a) + \Delta a\ \xi$$
$$W(\eta) = W(b) + \Delta b\ \eta,$$

where

$$(15) \qquad \begin{cases} \Delta a = W(\xi_0) - W(a) \\ \Delta b = W(\eta_0) - W(b), \end{cases}$$

half of the increment of W within a zone. Then we have

$$(16) \quad B_{ab} = \int_{-\xi_0}^{\xi_0} \int_{-\eta_0}^{\eta_0} \cos \left\{ k [W(\xi) - W(\eta)] \right\} d\xi\, d\eta = A_{ab} \cos \left\{ k [W(a) - W(b)] \right\} S(\Delta a)\, S(\Delta b),$$

and for the elements on the x-axis

$$(17) \quad B'_{ab} = A_{ab} \cos \left\{ k [W(a) - W(b)] \right\} \times \left\{ [S'(\Delta a)]^2 + [S'(\Delta b)]^2 + 2S'(\Delta a) S'(\Delta b) \times \cos \left[\frac{k}{2} (\Delta a + \Delta b) \right] \right\},$$

where

$$(18) \qquad \begin{cases} S(\Delta) = \dfrac{\sin k\Delta}{k\Delta} \\ S'(\Delta) = \dfrac{1}{2} \dfrac{\sin (k\Delta/2)}{(k\Delta/2)}. \end{cases}$$

When $\Delta a = \Delta b$ or $\Delta a \doteqdot \Delta b$, we can put $B_{ab} = B'_{ab}$. Δa, Δb are obtained from the wave aberration curve, and (18) can be calculated numerically.

Thus, for given ω, $D(\omega)$ can be got by summing the products A_{ab}/π, $\cos\left\{k[W(a) - W(b)]\right\}$, $S(\Delta a)$ and $S(\Delta b)$ for various (a, b).

In the experiments with the interferometer, the incoherent illumination is obtained by opening the iris stop I (fig. 1) fully and by inserting ground glass between condenser and object.

In figure 7 (c) and figure 8 (c), the response function and the image in incoherent illumination are shown, where the response function is translated by (11) into the response to a square topped pattern $D'(\omega)$. They are obtained at the same focal position as the case of (b).

At the first zero of $D'(\omega)$, where $\omega_1 = 0.45$ ($R_1 = 540$) the contrast of pattern vanishes and this is the resolution limit defined previously. Between this and the second zero $\omega_2 = 0.65$ ($B_2 = 770$), $D'(\omega)$ is negative, accordingly contrast reversal as is shown in figure 8 (c). Beyond ω_2 contrast returns to the former state. Between the third and the fourth zero, ω_3 and ω_4, $D'(\omega)$ take the maximum value at $\omega_f = 1.5$ ($R_f = 1800$).

Figure 9 (c) and figure 10 (c) show the response functions and the image of the object in incoherent illumination. In Table 1, the results in this case are also shown.

These results show that the contrast of image becomes smaller as frequency increases and the resolution limit is extended to double the frequency given by central coherent illumination, showing the agreement with ABBE's theory.

In this case it is also clear that the monofrequency grating is not a suitable object to determine the resolution limit.

5. Conclusion. — The correspondence between the aberration and the image definition of the microscope objectives were studied with the interferometer specially designed for this purpose.

In the experiments, the image observed is formed by light passing through the objective twice (to and fro) reflected by M_2. However, as the results of the experiments show good agreement with calculation, the definition of the image and the resolving power in the actual case, in which ray passes only once through the objective, can be estimated from its wave aberration in the same way as was shown in this paper.

The importance of the coherence of the illumination pointed out by HOPKINS was actually shown in these experiments by comparing perfectly coherent and incoherent illumination. But in the practical case, the illumination is neither coherent nor incoherent, but partially coherent. The results in such cases are being investigated.

Finally, the authors wish to express their sincere thanks to Prof. H. KUBOTA for his encouragement.

REFERENCES

[1] R. V. SHACK, *J. Research Natl. Bur. Standards*, **56**, 1956, p. 245.
[2] F. TWYMAN, *Trans. Opt. Soc.*, Lond., **24**, 1922-3, p. 189.
[3] F. TWYMAN, *Prism and Lens Making*, Hilger and Watts ed., 1950, p. 440.
[4] Y. UKITA and J. TSUJIUCHI, *J. Mechanical Laboratory*, **8**, 1954, p. 37. (in Japanese).
[5] H. H. HOPKINS, *Wave Theory of Aberrations*, Oxford, 1950, p. 48.
[6] M. HERZBERGER, *J. O. S. A.*, **37**, 1947, p. 485.
[7] H. H. HOPKINS, *ibid.*, p. 21.
[8] P. M. DUFFIEUX, *L'integrale de Fourier et ses applications à l'optique*, Rennes, 1946.
[9] H. H. HOPKINS, *Proc. Roy. Soc. A.*, **217**, 1953, p. 408.
[10] J. W. COLTMAN, *J. O. S. A.*, **44**, 1954, p. 468.
[11] H. H. HOPKINS, *Proc. Roy. Soc. A.*, **231**, 1955, p. 91.
[12] M. DE, *Proc. Roy. Soc. A.*, **233**, 1955, p. 91.
[13] W. H. STEEL, *Rev. d'optique*, **32**, 1953, p. 143.
[14] Y. UKITA and J. TSUJIUCHI, *J. Mechanical Laboratory* of Japan, **2**, 1956, p. 1.

Manuscrit reçu le 15 novembre 1956.

Some methods for measurements on thin films

A. C. S. van Heel and A. Walther
National Research Council, Ottawa

Summary. — *Four modifications of* Françon's *interference eyepiece are given which enable one to measure the step in a wave front and at the same time the light loss. As it is possible to carry out measurements while depositing the layer, the methods described can be used for monitoring purposes, the non-optical windows not doing much harm to the appearance of the field of view. With the available apparatus a precision in the determination of the light path of* $\lambda/200$ *is easily obtainable for all layer thicknesses.*

Sommaire. — *Une description est donnée de quatre usages différents de l'oculaire interférentiel de* Françon *pour la mesure de l'épaisseur et des pertes de lumière produites par des couches minces. Les mesures peuvent s'effectuer dans le vide, l'influence des fenêtres non-optiques étant petite. Le procédé peut aussi servir de contrôle pendant l'évaporation. Avec les appareils en service les chemins optiques sont déterminés avec une précision* $\lambda/200$, *valeur indépendante de la grandeur du chemin optique.*

Zusammenfassung. — *Mit Hilfe des Interferenzokulares von* Françon *wird nach vier verschiedenen Verfahren die Phasenverschiebung und die Intensitätsschwächung einer ebenen Welle bei dem Durchgang durch eine dünne Schicht gemessen. Man kann die Messungen während der Herstellung der Schichten vornehmen, weil die das Vakuum abschliessenden Planfenster trotz ihrer mangelhaften optischen Güte das Bild nicht wesentlich beeinträchtigen. Mit den gebräuchlichen Hilfsmitteln kann man leicht eine Genauigkeit der Weglängenunterschiede von* $\lambda/200$ *erreichen ; sie ist unabhängig von der Schichtdicke.*

1. Introduction. — A phase object in the shape of a step, obtained by evaporating a thin layer on only one half of an object glass, modifies an incident plane wavefront in two ways : the layer causes both a phase retardation and a reduction in amplitude of one side of the wavefront with respect to the other. If it were possible to determine accurately this phase shift and transmission loss, one would be able to calculate both the thickness and the refractive index of the layer.

Further it would be useful to be able to measure these quantities during the evaporation. The first author has described a modification of the Françon interference eyepiece [1, 2, 3] which enables observations of this kind to be made when the object is situated within the bell jar [4]. This method was based on colour matching. Although some useful results were obtained, colour matching has the obvious disadvantage that the results are highly affected by the dependence of the transmission on the wavelength. This limits the precision to $\lambda/40$ at the most, or about 150 Å.

A great improvement can be obtained by using monochromatic light. In this paper some investigations in that direction are described. Some preliminary results with the monochromatic method have been given in reference 4.

2. Principle of the Françon eyepiece. — A linearly polarized plane wavefront after being deformed by the object passes through a Savart plate placed in 45° position. The wavefront is split into two mutually perpendicularly polarized wavefronts I and II, of equal amplitude. The phase difference between these two wavefronts depends linearly on the angle of incidence on the Savart plate ; moreover, I and II receive a lateral shift with respect to each other, which only depends on the thickness (and the material) of the Savart plate. After this splitting of the wavefront the light passes through an analyser ; the two wavefronts interfere and the interference is observed in the plane of the object.

Figure 1 shows the parts of the apparatus which are important for an understanding of Françon's method. For a more complete description of the arrangement within the bell jar we refer to [4].

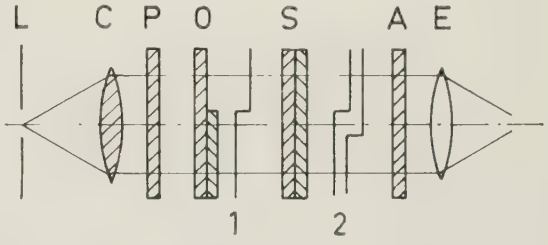

Fig. 1. — L : slit. C : collimator lens. P : polarizer. O : object. S : Savart plate. A : analyser. E. eyepiece. 1 and 2 : the wavefronts before and after passing through the Savart plate.

3. Application of the principle. — As mentioned in § 1, the object is assumed to be a step, obtained by evaporating a thin layer on one half of an object glass. The two wavefronts emerging from the Savart plate are shown in figure 2.

The parts C_1D_1 and C_2D_2 of the wavefronts have passed both substrate and layer ; their amplitude will be reduced by a factor $(1 - \rho)$ with respect to the amplitude of the wavefronts A_1B_1 and A_2B_2, which have passed only through the substrate. The square of $(1 - \rho)$ is the ratio of the transmission through substrate plus layer to that through the substrate alone. α is the phase difference introduced by the Savart plate. The latter may be varied by changing the angle of incidence of the light on the Savart plate. This can be done either by rotating the Savart plate itself, or with the aid of an auxiliary rotating plane parallel plate in a divergent part of the incident pencil. When this plate is rotated, φ changes because of a shift of the virtual image of the slit [4].

We assume that the analyser is parallel or crossed with respect to the polarizer. Using monochromatic light, we will in general observe three different intensities in the three parts of the field of view, formed by the strip and the background on either side (p, q and r, in figure 2). In white light, one would see in general three different colours.

The problem is to measure accurately the quantities α and ρ. From these quantities it is possible to derive the refractive index and the optical thickness of the layer.

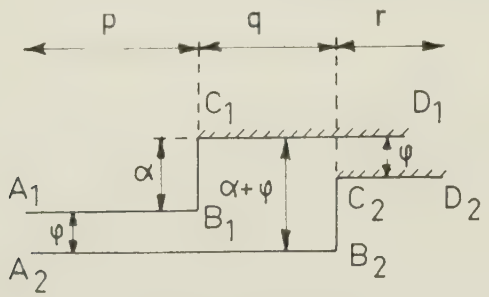

Fig. 2. — All the angles in the figure are supposed to be positive.

4. Methods of measurement. — In this section we will deal with four different methods of measurement in monochromatic light. Technical details will be discussed later.

Method 1. — There exist values of φ, spaced at intervals of 2π, for which the intensities in the left- and righthand sides of the field of view are zero, and also a similar series of φ values, for which the intensity in the strip is a minimum. α can be determined (with an uncertainty of $2k\pi$, k integral) from the relative displacement of these two series.

Using this method, the possible error in α expressed in wavelengths will never be much smaller than $\lambda/60$, and this can only be reached by repetition of measurements. In addition ρ can not be measured by this method. Although setting on minima is not a good criterion to base a measurement on, this method appears to be more reliable than colour matching.

Better than this rather obvious method is :

Method 2. — Let us first assume that the transmission loss ρ is zero. Then the intensities in the left- and righthand side of the field of view are always the same, and it is possible to choose φ in such a way, that the strip intensity is matched with the field intensity. Denoting the interfering amplitudes by a, we get for this particular case the relation :

$$(1) \quad a^2 + a^2 + 2a^2 \cos\varphi = a^2 + a^2 + 2a^2 \cos(\varphi + \alpha)$$

from which it follows, that

$$(2) \qquad \varphi = -\frac{1}{2}\alpha + k\pi,$$

provided that α is not equal to $2k'\pi$.

At successive matching points φ changes by an amount π, and the intensity is alternately high and low. We see that for this method it is necessary to determine the zero point of calibration, that is to say the positions of the Savart plate (or of the auxiliary plane parallel plate) in which the phase difference φ equals $2k\pi$. This is a disadvantage. However we will show now that with this method of matching it is possible to measure the transmission loss.

In general ρ is not zero, and then it is impossible to match the intensities of the three parts of the field of view at the same time. But one can match either the intensities of the lefthand side of the field of view and the strip ($\varphi = \varphi_1$), or those of the righthand side and the strip ($\varphi = \varphi_2$).

The resulting relations are

$$(3) \quad a^2 + a^2 + 2a^2 \cos\varphi_1 = a^2 + (1-\rho)^2 a^2 + 2a^2(1-\rho)\cos(\varphi_1 + \alpha)$$

and

$$(4) \quad a^2 + (1-\rho)^2 a^2 + 2a^2(1-\rho)\cos(\varphi_2 + \alpha) = a^2(1-\rho)^2 + a^2(1-\rho)^2 + 2a^2(1-\rho)^2 \cos\varphi_2.$$

This can be reduced to

$$(5) \quad 2(1-\rho)\cos(\varphi_1 + \alpha) = 2\cos\varphi_1 + (2\rho - \rho^2)$$

and

$$(6) \quad 2(1-\rho')\cos(\varphi_2 + \alpha) = 2\cos\varphi_2 + (2\rho' - \rho'^2)$$

in which ρ' is defined by

$$(7) \qquad \rho' = \frac{-\rho}{1-\rho}.$$

In a range of 2π there are two values for φ_1 and two values for φ_2 that satisfy equations (5) and (6) respectively. In general one pair corresponds to a higher, and one pair to a lower intensity. As there are only two unknown quantities, α and ρ, the problem is overdetermined. This appears to be useful, not only for a check, but also for calculating a final small correction to the zero point of calibration.

The accuracy of this method depends largely on the value of α to be measured. We have tested the method on a layer of zinc sulphide of an optical thickness of about 0.7λ; the precision in this case appeared to be about $\lambda/125$ in α and 2 % in the energy transmission $(1-\rho)^2$. This precision could again only be obtained by repetition of measurements.

Since the extraction of α and ρ from the equations (5) and (6) is an impractically long calculation, this method is not useful in practice, and we shall not discuss the required precision of adjustment.

Method 3. — The first method that we have dealt with suggests the use of some half shade analyser. This is not quite as easy to apply as it seems, because the light emerging from the Savart plate is, in general, elliptically polarized, and so an ordinary half shade device with a half shade angle of a few degrees will be of no use, as will be shown.

We require an analyser, consisting of two parts, whose directions of polarization are different, and whose line of separation is perpendicular to the strip.

Let this line of separation be horizontal, the strip being vertical.

Let the amplitudes of the mutually perpendicular vibrations that emerge from the SAVART plate be a and b ($a \geqslant b$). Let the angle between the a-vibration and the bisector of the two directions of vibration of the analyser be θ. The half shade angle is 2β. Let ψ be the phase difference between a and b (fig. 3).

We stress, that a, b and ψ will have different values in the three parts of the field of view. In the left and right hand side of the field of view a is equal to b and ψ is the angle φ of section 3. In the

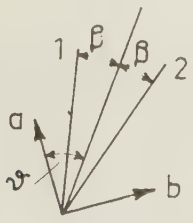

FIG. 3. — 1 and 2 : directions of polarization in the halfshade device.
N. B. : a, b and the phase difference ψ have different values in the three parts of the field of view (see text).

strip ψ is the angle $\varphi + \alpha$ and a is in general not equal to b.

The condition for matching one of the upper parts of the field of view with its lower part is

$$a^2 \cos^2(\theta + \beta) + b^2 \sin^2(\theta + \beta) + 2ab \sin(\theta + \beta) \cos(\theta+\beta) \cos\psi = a^2 \cos^2(\theta - \beta) + b^2 \sin^2(\theta - \beta) + 2ab \sin(\theta - \beta) \cos(\theta - \beta) \cos\psi, \quad (8)$$

or

$$\cos\psi = \frac{a^2 - b^2}{2ab} \operatorname{tg}^2\theta. \quad (9)$$

If we make $\theta = 0$, the condition for matching becomes

$$\cos\psi = 0. \quad (10)$$

That is $\psi = \pi/2 + k\pi$, when the upper and lower parts are matched, and this condition is independent of ρ. We observe, that it is even independent of the adjustment of the polarizer with respect to the SAVART plate (1).

We will prove now that β determines the precision of the settings. The contrast V between the upper and lower halves of the field of view with intensities J_1 and J_2 can be defined as

$$V = \frac{J_1 - J_2}{J_1 + J_2} = \quad (11)$$

$$= \frac{(a^2 - b^2)\sin 2\theta \sin 2\beta - 2ab \cos 2\theta \sin 2\beta \cos\psi}{(a^2+b^2)+(a^2-b^2)\cos 2\theta \cos 2\beta + 2ab \sin 2\theta \cos 2\beta \cos\psi}.$$

The matching condition requires θ to be zero. Assuming that the smallest perceptible $V = 0.01$, we see, that

$$|\cos\psi| \approx |\Delta\psi| < 0.01 \times \times \frac{(1 + b^2/a^2) + (1 - b^2/a^2)\cos 2\beta}{2(b/a)\sin 2\beta}. \quad (12)$$

b/a varies in practice between 0.7 and 1.0, and so the possible error in ψ is smallest, when 2β is about 90° ; it might even be made slightly greater. Then the error in ψ is of the order of magnitude of 0.01 rad, or about $\lambda/600$.

An important point is the precision required in the adjustment of θ. Since we can ensure that $|\cos\psi| < 0.01$, the matching condition (9) requires

$$|\operatorname{tg} 2\theta| < 0.01 \frac{2ab}{a^2 - b^2}. \quad (13)$$

If b/a is never smaller than 0.7, we must ensure that

$$|\theta| < 0.014. \quad (14)$$

if an accuracy of $\lambda/600$ is to be obtained. This is feasible.

This method appears to be very useful. However, it is impossible to get information about transmission losses in this way. Also it is very difficult to manufacture a half shade device with a narrow line of separation.

With a provisional half shade analyser, composed of two small pieces of Polaroid sheet, the reproducibility of readings appeared to be of the order of magnidute of $\lambda/150$. It is clear that this might be considerably improved by a better construction of the half shade device.

Method 4. — It is well known that a SAVART plate, used in conjunction with an analyzer in the 45° position, is one of the most delicate means of detecting a small quantity of polarized light in a beam consisting of mainly unpolarized light [5]. It can easily be seen that this « SAVART Polariscope » is also a very sensitive instrument for distinguishing linearly polarized light from elliptically polarized light.

Let a convergent beam, consisting of elliptically polarized light, be incident on a SAVART plate. Then the elliptical vibration can always be decomposed into two linear vibrations along the directions of vibration of the SAVART plate, and it is certain that neither of these components is zero. Consequently, the interference pattern after the light has passed through an analyzer, will consist of straight fringes at infinity, no matter how the SAVART plate is orientated with respect to the ellipse. The sensitivity, viz. the visibility of the fringes, may be increased by turning the analyzer to an appropriate position. The attainment of the best visibility might be hampered by insufficient brightness of the light source or imperfections in the SAVART plate and the polaroid.

(1) For $\psi = 45^\circ$ matching is possible only when $a = b$. This property might be used for the adjustment of the polarizer with respect to the analyzer. We do not need to do this when we use this third method of measurement. But it is useful to adjust the apparatus for methods 1 and 2.

However, if the incident pencil of light is linearly polarized, there are two positions of the Savart plate in which the direction of the incident vibration coincides with one of the directions of vibration of the Savart plate. In such a case there is only one emergent vibration and so there are no interference fringes for any position of the analyzer.

We see then, that vanishing of fringes indicates not only the fact that the incident lightwave is linearly polarized ; also the direction of that vibration can be determined (²). As mentioned before, the sensitivity is only limited by the brightness of the lightsource and the imperfections of the Savart plate and the analyser.

It can be easily shown how this property can be used for our measurements. An elliptically polarized wave emerges from the Savart plate S_0 of the Françon device. It will only degenerate into a plane polarized wave in the field (or strip) if φ (or $\varphi + \alpha$) equals $k\pi$. A second Savart plate (denoted by S_1) can now be used as an indicator of phase differences $k\,\pi$. Moreover, the direction of the linear vibration that we obtain in this way depends on the ratio of the amplitudes of the composing vibrations. As we are also able to measure this direction, the transmission loss ρ can be measured too.

To set up the apparatus for this method, we remove the analyzer in the apparatus of figure 1. A second Savart plate S_1 followed by an analyzer are now put just behind the eyepiece. We now change φ and the orientation S_0 of S_1 simultaneously. Settings are made in which the fringes formed by S_1 disappear, first in the strip, and then in the remainder of the field of view. The value of α, with an uncertainty of $k\pi$, follows immediately from the readings of φ. Four settings are made, two in the strip and two in the remainder of the field of view, each pair with a φ-difference of π. When we use four adjacent positions of the second Savart plate S_1, denoted by ζ_{os} $\zeta_{\pi s}$ ζ_{of} and $\zeta_{\pi f}$, where the subscripts s and f refer to strip and field, these four angles are related as follows : $\zeta_{os} + \zeta_{\pi s} = \zeta_{of} + \zeta_{\pi f}$ and it can readily be shown, that the light loss satisfies the equation

$$1 - \rho \quad \frac{\operatorname{tg}\left\{45 \pm \frac{1}{2}\,(\zeta_{\pi f} - \zeta_{of})\right\}}{\operatorname{tg}\left\{45 \pm (\zeta_{\pi f} - \zeta_{os})\right\}}\,.$$

The sign must be chosen so that $\rho \leqslant 1$.

When the polarizer is placed in the 45° position with respect to S_0, ζ_{of} is equal to $\zeta_{\pi f}$. But we want to stress that this adjustment is not necessary ; the position of the polarizer and the analyser is not relevant to the measurements.

One problem remains ; α has been determined only with an uncertainty of $k\pi$. In the former three methods this was a serious difficulty, here however it is easily solved [4]. One rotates S_1, until the fringes are perpendicular to the strip. Then the monochromatic lightsource is replaced by a white one. The coloured fringes obtained in this way are seen to be shifted in the strip with respect to those in the remainder of the field. The shift is a direct measure of α, and can be estimated with a precision of at least $\lambda/10$. This is sufficient. Experiments have shown that with Savart plates and polaroids of rather poor quality a precision of $\lambda/200$ on each setting for φ ,and of 2 % for the transmission factor $(1 - \rho)^2$ could be obtained without any difficulty.

Since during the construction of this apparatus no careful adjustment of the optical parts responsible for interference is required, and since α and ρ are separately determined from different readings, this measuring device gives reliable results and is easy to handle.

5. **Some technical details.** — No matter which method we use, we have to surmount some technical difficulties.

Firstly, special care has to be taken in the manufacture of the Savart plates. The flatness of the surfaces of each of their components is fairly important, but it is still more important to use plates of as uniform a thickness as possible. Again, in a Savart plate the two component plates are so rotated with respect to each other, that their principal sections are at right angles. Both to obtain this as exactly as possible, and to avoid additional double refraction caused by mechanical stresses, we recommend that the Savart plates (at least S_0, but if possible also S_1) be mounted uncemented. Equal thickness of the two components is of minor importance, as a small difference in thickness gives only rise to a shift of the fringes at infinity. Inhomogeneities of the material are hard to avoid; they should be as small as possible.

The slit width is limited by the requirement that the angular width of the parallel pencil incident on S_0 should not be larger than one twentieth of the angular fringe distance. This suggests the use of a thin S_0 plate, but such a plate would produce only a small lateral shift. A compromise must be made.

Polaroid sheet, whether mounted between glassplates or not, is rather grainy and the degree of polarization it yields is often not high enough for our purpose. Moreover it is often double refracting because of mechanical stresses. For measurements in which the greatest possible precision is required, we recommend the use of Nicol prisms.

During our experiments we have had a considerable amount of trouble because of double refraction of the lenses and of the substrate on which the layer is deposited. One might place the lenses as far as possible between parallel Polaroids, but in that case one obtains intensity variations in the wavefront that make the measurements less reliable.

(²) This second application is used in the Wild « polaristrobometer » (See f. i. in Glazebrook's Dict. of Appl. Physics IV, p. 480 ; London, Mac Millan, 1923).

6. **Some results.** — In our final experimental apparatus we used the green mercury line (5 461 Å) of a super high pressure mercury arc. The line was isolated using a Kodak-Wratten filter together with a green gelatine filter to cut out the remaining amount of red light. The focal length of the collimator lens was 1 m. The slit width was a few tenths of a millimeter, the slit height about 30 mm. The filter was placed before the slit. The polarizer (Polaroid sheet mounted between glass plates) was placed close after the collimator lens. Then came the object ; a step of zinc sulphide, evaporated on a rather good optically worked plate of plane parallel glass. Next came the SAVART plate S_0 (composed of two plates of calcite each of two millimeters thickness). It was mounted on a WILD theodolite T-2, with which could be measured with a precision of one tenthousandth of a degree. The angular distance between two successive fringes of S_0 was about 0.002 rad, it produced a lateral shift of about 0.3 mm. The SAVART plate was followed by an analyzer (same as the polarizer), followed by a microscope with a magnification of about fifteen times, and a free working distance of about 10 cm. The microscope was focussed on the object. In method 4 the analyzer was removed and replaced by a SAVART S_1 (quartz, thickness two times 13 mm) mounted in a divided circle followed by an unmounted piece of polaroid sheet, both placed just after the eyepiece.

The results given below are averages of series consisting of five to ten measurements. The results obtained for α were reduced to path differences in air.

A. Results obtained with methods 1 and 2. — It is clear that methods 1 and 2 should both be used at the same time, as one has to use method 1 to find the zeropoint of calibration for method 2 anyway.
For one object the results were

Method 1 : $\alpha = 0.181 \pm 0.01\ \lambda$,
Method 2 : $\alpha = 0.167 \pm 0.01\ \lambda$, $(1-\rho)^2 = 77.4 \pm 2\%$.

By putting two objects in the apparatus at the same time, it is possible to determine both the sum and the difference of their α values, and ratio and product of their transmissions $(1-\rho)^2$. This can be done by placing the two objects in the appropriate position with respect to each other. The difference of the values of α and ρ for a second object and the one mentioned above (both of which were made at the same time) was determined ; it appeared to be smaller than our limit of accuracy.

For the sum of these two objects we found

Method 1 : $\alpha = 0.351 \pm 0.01\ \lambda$,
Method 2 : $\alpha = 0.375 \pm 0.01\ \lambda$, $(1-\rho)^2 = 62.4 \pm 2\%$.

This would give for one layer (the two objects being identical within our limit of accuracy) :

Method 1 : $\alpha = 0.176\ \lambda$,
Method 2 : $\alpha = 0.188\ \lambda$, $(1-\rho)^2 = 79.0\%$.

These results are not particularly good. Two explanations might be given. Firstly : the three parts of the field of view should have a perfectly uniform intensity. However it was impossible to obtain this because of the rather poor quality of both the Savart plate and the polarizer and analyzer. Secondly, the tested objects were rather old and had been used many times, with the result that they were more or less damaged. This did not favour a uniform field either.

For testing method 4 new objects were prepared.

B. Results obtained with method 4. — Firstly the values of α for two new layers were determined.

First layer : $\alpha = 0.6994 \pm 0.0045\ \lambda$,
Second layer : $\alpha = 1.442 \pm 0.006\ \lambda$.
This leads to a sum of $2.141 \pm 0.007\ \lambda$.
Then the sum was actually measured, the result being

$$2.137 \pm 0.007\ \lambda.$$

Considering that the uniformity of the intensity in the field of view was still rather poor, this agreement is very good. Care was taken always to take measurements on the same part of the objects. Next the phase difference α and the transmission were measured simultaneously at another spot in one of the layers. The results were

$$\alpha = 1.411 \pm 0.006\ \lambda,\ (1-\rho)^2 = 93.9 \pm 2\%.$$

Now as mentioned in the introduction, it is possible to calculate from these data the thickness and the refractive index of the layer. This can be reversed, too ; assuming the refractive index to be known, the transmission can be calculated from the value of α (assuming that there is no absorption). In this paper we will not describe the formulae involved. We will only mention that they can be easily derived from the well known theory of thin films (see e. g. ref. 6, in which several references are listed). Taking the refractive index to be 2.40, the value of the transmission as derived from the phase difference was $93.8 \pm 2\%$, a very satisfactory result.

7. — In the introduction we made the statement, that for our purpose the method of colour matching [7] is not sufficiently accurate. In this last section we will describe one of the experiments we performed to arrive at this conclusion.

We used the arrangement of figure 1, provided with a white lightsource. As an object we used a ZnS film of a geometrical thickness of roughly 5550 Å. A crude calculation gives for the transmission of this film, relative to the glass substrate : maxima of 100 % at wavelengths of 4400 Å, 5300 Å and 6600 Å, and minima of about 70 % at 4800 Å and 5900 Å.

We measured the angle over which the SAVART plate had to be rotated to go from the sensitive purple on the " lefthand side " of the central white to the same colour on the " righthand side " of the central white. We did this twice ; first in the part of the field of view in which the wavefronts were unaffected by the ZnS film, and then in the strip, i. e. in the region in which one of the two wavefronts was affected by the ZnS film, and the other wavefront was not.

The two measured angles differed by 5 %, whereas the standard error in the individual measurements was about 0.5 %. Large deviations like the one found here cannot be tolerated.

Considerations of this kind made us investigate the possibilities of the monochromatic methods.

Acknowledgements. — We sincerely wish to express our grateful thanks to the National Research Council of Canada, and especially to the Director of the Applied Physics Division Dr L. E. HOWLETT, who by inviting us to this laboratory made these investigations possible, and who was always ready to meet our special wishes.

REFERENCES

[1] M. FRANÇON, *Rev. Opt.*, **31**, 1952, 65.
[2] M. FRANÇON, *Rev. Opt.*, **32**, 1953, 349.
[3] M. FRANÇON, *Opt. Acta*, **1**, 1954, 50.
[4] A. C. S. VAN HEEL, *Can. J. Phys.*, **34**, 1956, 1460.
[5] A. SCHUSTER, *An introduction to the theory of optics*, London, Arnold & Co, p. 213, 3d ed., 1928.
[6] O. S. HEAVENS, *Optical properties of thin solid films*, London, Butterworth Sc. Publ., 1955, chapter 4.
[7] E. INGELSTAM, I. C. O. Conference, Boston, 1956.

Manuscrit reçu le 16 décembre 1956.

Einige Beobachtungen über die Wahrnehmbarkeit kurzdauernder, farbiger Lichterscheinungen und ihre Deutung *

O. STIERSTADT und W. ZWIESLER

ZUSAMMENFASSUNG. — *Es wurden kurzdauernde, verschiedenfarbige, an der Grenze der Wahrnehmbarkeit liegende Lichtsignale mit gemischt foveal-extrafovealer Beobachtungsweise untersucht. Unter Ansatz einer neuen Definition für die Schwellenwerte ergaben die Beobachtungen auch in diesem Falle die Gültigkeit des Gesetzes von* BLONDEL *und* REY.

Die Bestimmung der einfachen und spezifischen Schwellenwerte führte zu dem überraschenden Ergebnis, dass unter gleichen Voraussetzungen grünes Signallicht einen rund 7 mal niedrigeren Schwellenwert liefert als rotes. Theoretische Betrachtungen führten zur Erklärung dieses Phänomens, dessen Ursache in der gemischt foveal-extrafovealen Beobachtungsweise zu suchen ist.

Im Hinblick auf das Problem der Sicherheit im Verkehrswesen dürfte dieses Resultat recht bedeutungsvoll sein. Weitere Untersuchungen zu dieser Frage z. B. durch eine Fortsetzung der Messungen im Freien unter verschiedenen atmosphärischen Bedingungen erscheinen uns als geboten.

SUMMARY. — *Light signals of short duration at the limit of perception and of different colours have been observed simultaneously both foveally and extrafoveally. For a given definition of the threshold value, measurements confirm the law of* BLONDEL *and* REY.

Determination of the specific threshold values leads to the surprising conclusion that, under corresponding conditions, a green signal gives a threshold about seven times lower than a red one. This phenomenon can be explained by the simultaneous foveal and extrafoveal vision. These conclusions are of importance in considering traffic control problems ; further observations, for example in the open air under different atmospheric conditions seem desirable.

SOMMAIRE. — *Des signaux lumineux de courte durée et de couleurs différentes, choisis à la limite de perception, sont observés à la fois en vision fovéale et extra-fovéale. Pour une définition donnée de la valeur du seuil, les mesures montrent dans ce cas que la loi de* BLONDEL *et* REY *est valable.*

La détermination des valeurs spécifiques du seuil amène à cette conclusion surprenante que, dans de mêmes conditions, un signal lumineux vert donne une valeur du seuil environ sept fois inférieure à celle du signal rouge. Ce phénomène peut être expliqué par la vision simultanée fovéale et extra-fovéale.

Il devrait en être tenu compte dans les problèmes concernant la sécurité de la circulation ; des observations seront effectuées à cette fin, en plein air, dans différentes conditions atmosphériques.

1. **Vorbemerkungen.** — Für die Beurteilung der Wahrnehmbarkeit kurzdauernder Lichtsignale ist ihre Erkennbarkeit bei Dunkelheit als der in der Praxis besonders interessante Fall wichtig.

Bekanntlich haben A. BLONDEL und J. REY [1-4] für schwache, an der Schwelle der Wahrnehmbarkeit liegende « weisse » Lichtblitze von kurzer Dauer das folgende Gesetz aufgestellt :

$$(B - B_0) \, . \, t = \alpha \, . \, B_0 \, ,$$

worin B_0 die Beleuchtungsstärke am Auge des Beobachters für $t = \infty$ und α = const. ist. Sie bestimmten α zu Werten zwischen 0,15 und 0,30 sec. Sie beobachteten ferner, dass dieses Gesetz bei sehr kurzen Lichtblitzzeiten, nämlich im Bereich von 1-100 msec, in die Form

$$B \, . \, t \quad \text{const.}$$

übergeht. Erst von etwa 0,1 sec ab bis zu ca. 1 sec Dauer der Lichtblitze gilt die erweiterte Gleichung. Dieses vereinfachte Gesetz konnte auch von anderen Autoren [5-7] bei sehr kurzen Zeiten bestätigt werden. Jedoch ergaben sich auch Abweichungen und Spezifikationen [7 ff.].

Demgegenüber ergaben frühere Untersuchungen des Forschungsinstituts für Physik [11] im Rahmen von FREILAND-Versuchen eine so grosse Streuung der Werte, dass für dieses Zeitintervall keine eindeutige Entscheidung zugunsten einer der beiden Formeln möglich war. Auch H. PIÉRON [12,13] hatte schon früher andere Ergebnisse erhalten. Von Belang sind in diesem Zusammenhang noch die Arbeiten von RICCO, CHARPENTIER, ASHER und LOESER [14-17]. Sie fanden, dass bei einem Sehwinkel bis zu etwa 1° im fovealen Sehen nur der auf das Auge fallende

(*) Mitteilung aus dem Forschungsinstitut für Physik, Heidenheim/Brenz.

Unserer Mitarbeiterin Dipl. Ing. ERIKA HELLWIG danken wir für ihre Mithilfe bei der Durchführung der vorliegenden Arbeit.

Dem US Air Research and Development Command (ARDC), European Office, Brüssel, danken wir für die Unterstützung dieser Arbeit durch Bereitstellung der Mittel. Die Arbeit wurde nach unseren Vorschlägen für ARDC, USAF, durchgeführt.

Lichtstrom für den Schwellenwert massgebend ist, nicht aber die Grösse der leuchtenden Fläche.

Während sich der Ausdruck « Schwellenwert » (auch « einfacher Schwellenwert ») auf « weisse » Lichterscheinungen bezieht, spricht man bei farbigen Lichtsignalen vom « spezifischen Schwellenwert ». Er ist für die einzelnen Farben unterschiedlich, da die Empfindlichkeit der verschiedenen Zäpfchenarten (für rot, grün und blau) ebenfalls variiert. Der einfache und der spezifische Schwellenwert sollen im allgemeinen durch ein « farbloses Intervall » getrennt sein, in dem man den Lichtreiz als solchen zwar bereits wahrnimmt, jedoch noch nicht spezifisch als Farbe. Indessen ist dies nicht immer bestätigt worden [18 ff.]. Andererseits wird die Farbewahrnehmung beim fovealen Sehen, wie König [22] feststellte, mit Ausnahme von gelb stets zugleich mit der Lichtwahrnehmung erkannt. Dies ist bemerkenswert, wenn man berücksichtigt, dass das Auge in dem aus dem kontinuierlichen Spektrum bestehenden Lichtgemisch stets gerade auf die Farbe gelb akkomodiert [23]. Auch v. Kries, Nagel und Lummer [24, 25] konnten den Nachweis führen, dass ein farbloses Intervall beim fovealen Sehen nicht auftritt. Zu ähnlichen Ergebnissen kamen Rouse [26] und Davy [27].

Schliesslich ist in diesem Zusammenhang noch auf die Arbeiten von Broca und Sulzer hinzuweisen [28-30], sowie auf diejenigen von N. E. G. Hill [36], in denen ähnliche Probleme wie in der vorliegenden Arbeit behandelt werden, allerdings unter anderen Voraussetzungen und Versuchsbedingungen.

2. **Versuchsbedingungen.** — Die Wahrnehmbarkeit einer Lichterscheinung kann nicht als eine Funktion einer einzigen Veränderlichen dargestellt werden. Es ist vielmehr notwendig, diejenigen Komponenten für die Untersuchungen auszuwählen, die je nach den in der Praxis gegebenen Verhältnissen von Belang sind.

a. Von uns wurden kurzdauernde, an der Grenze der Wahrnehmbarkeit liegende Lichtblitze verwendet. Die Lichtblitzdauer wurde auf den Bereich von 2-200 msec beschränkt, weil die Empfindlichkeit des Auges offenbar nur in diesem Bereich wesentlich variiert.

b. Es wurden Lichtquellen benutzt, die unter einem Sehwinkel von weniger als 1° erschienen, entsprechend dem Umstand, dass diese Arbeit sich auf die Wahrnehmbarkeit speziell von Signallichtquellen bezieht. Diese sollen schon in möglichst grosser Entfernung erkannt werden ; dabei werden sie im allgemeinen praktisch als punktförmig erscheinen. Dann kommt es unabhängig von den geometrischen Daten nur noch auf den Lichtstrom an.

c. Der zeitliche Abstand der einzelnen Lichtblitze wurde wesentlich grösser als 1,5 sec gehalten, damit sie nicht « überschwellig » werden konnten.

d. Es wurde mit dunkel-adaptiertem Auge ohne Fixationspunkt, jedoch bei geringer allgemeiner Raumhelligkeit gearbeitet. Dies entspricht den in der Praxis gegebenen Verhältnissen (z. B. Sternhimmel, Nachthelle, Grosstadthelle usw.), bei denen meist nur die ungefähre Richtung bekannt ist, in der sich die Signallichtquelle befindet. Die Adaptionszeit wurde auf ca. 20 min beschränkt. Die geringe Raumhelligkeit ermöglichte eine gewisse, Orientierung, so dass die Beobachtungsweise nur zum Teil extrafoveal erfolgte.

e. Zur Beobachtung der Lichtblitze wurden jeweils 8-10 Personen eingesetzt, farbentüchtige deren Alter zwischen 20 und 50 Jahren lag.

f. Da die Empfindlichkeit des Auges für die verschiedenen Spektralbereiche relativ zum Adaptionszustand variiert [31], wurden die Beobachtungen mit farbigen Lichtsignalen unter genau den gleichen Versuchsbedingungen wie diejenigen mit weissem Licht ausgeführt. Man kann die so ermittelten einfachen und spezifischen Schwellenwerte dann unmittelbar miteinander vergleichen.

g. Nach den aus dem Forschungsinstitut für Physik stammenden früheren Vorschlägen [11] wollen wir, abweichend von der üblichen Definition, diejenigen Lichtsignale als in der Schwelle der Wahrnehmbarkeit liegend bezeichnen, bei denen — über alle Versuchspersonen gemittelt — gerade die Hälfte der ausgestrahlten identischen Lichtblitze wahrgenommen wird. Man vermeidet mit dieser Definition die Abhängigkeit des Schwellenwertes von dem Sinne der Intensitätsänderung beim Durchlaufen der Schwellenintensität. Zufällige Beobachtungsfehler werden weitgehend ausgeschaltet und ausserdem die Erfordernisse der Praxis besser als mit der bisherigen Definition berücksichtigt.

3. **Versuchsanordnung.** — Es wurden Lichtblitzgeber konstruiert, die in der Abb. 1 schematisch dargestellt sind. Ihre nähere Beschreibung erübrigt

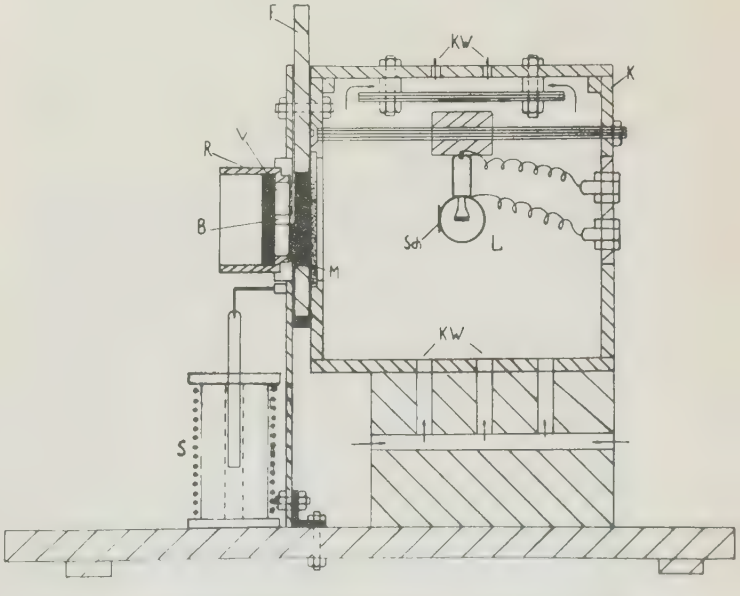

Abb. 1. — Lichtblitz-Geber, schematisch. Massstab 1 : 2,5.

sich (kleine Spezial-Projektionslampe L, axiale Lage der Wendel ; Schirm Sch, um die direkte Beleuchtung der Mattscheibe zu vermeiden ; Gehäuse mit Zinkweiss ausgestrichen ; Kalt-Warmluft-Zirkulation KW zum Temperaturausgleich). — Um möglichst punktförmige Lichtblitze zu erreichen, wurde die Blende auf eine Öffnung von 3 mm ⌀ eingestellt und aus einer Entfernung von 9-10 m beobachtet. Das entspricht einem Sehwinkel von etwa 1′ 15″. Nach HOOKE liegt die Grenze des Auflösungsvermögens des menschlichen Auges bei 1′. Womit die Versuchsbedingungen gemäss Abs. 2 b eingehalten sind.

Nach der schematischen Darstellung in Abbildung 1 besitzen die Lichtblitzgeber vor der Mattscheibe einen fotografischen Verschluss V (Compur-Verschluss), mit dem die Dauer der Lichtblitze in dem nach Abs. 2 a der Versuchsdingungen zugelassenen Bereich von 2-200 msec variiert werden kann.

Um bei den genannten Daten und Dimensionen zu Belichtungen zu gelangen, die im Bereich der Schwelle der Wahrnehmbarkeit liegen, mussten Schwächungsfilter vorgesetzt werden. Dazu dienten einmal das vorn am Verschluss eingeschraubte Rohrstück R (für feste Filter), zum anderen die Filterschienen F (für schnell veränderliche Filterkombinationen). Hier konnten (nach Art von Diapositiven) Filter genau bekannten Durchlässigkeitsmoduls eingesetzt werden. So war es möglich, durch Auswechseln der Schieber auch im Dunkeln schnell und sicher ganz bestimmte Intensitäten für die Lichtblitze einzustellen. Als Filter wurden hierzu Grauglasfilter der Type NG 5 von der Firma Schott & Gen. Mainz, in verschiedenen Stärken benutzt. Bei allen Beobachtungen wurde die Auswahl der Filterscheiben so vorgenommen, dass die Beobachter nicht durch einen bestimmten Gang der Intensitätssänderung beeinflusst werden konnten.

Unterhalb des Compurverschlusses ist eine Spule S angebracht, in die ein mit dem Verschluss gekoppelter Eisenstab eintaucht. Diese Einrichtung erlaubt es, den Verschluss auf elektromagnetischem Wege zu betätigen. — Für die Untersuchungen mit farbigem Licht wurden entsprechende Farbgläser in das Rohrstück R am Verschluss eingesetzt.

Bei der Ausführung der Versuche wurde bewusst von der üblichen Methode der Beobachtung mit einem besonders festgelegten Fixationspunkt abgewichen. Solche Fixations punkte sind in der Praxis meist auch nicht vorhanden. Vielmehr wurde durch eine sehr schwache allgemeine Helligkeit in dem Raum, in dem sich der Lichtblitzgeber befand, für eine Anpassung an die Verhältnisse in der Praxis gesorgt (z. B. Nachthimmel-Helligkeit). — Da das « Klicken » des Verschlusses bei jedem Lichtblitz von den Beobachtern gehört wird, wurde der Kunstgriff angewandt, ausser den sog. echten Lichtblitzen bei jeder Serie noch eine zusätzliche Anzahl « blinder » Blitze auszusenden, deren Anzahl und Reihenfolge den Beobachtern selbstverständlich unbekannt blieb. Bei den blinden Blitzen wurde lediglich durch ein Verdecken der Blendenöffnung dafür gesorgt, dass kein Blitz gesehen werden konnte. Eine an sich mögliche, auf reiner Einbildung einzelner Beobachter beruhende Beeinflussung der Beobachtungsergebnisse durch das Verschlussgeräusch wurde damit praktisch ausgeschaltet. — Ausserdem wurden unter Beibehaltung der Verdunkelung zwichen den einzelnen Versuchsserien Pausen eingeschoben, um den Beobachtern die Möglichkeit zu geben, die Augen ausruhen zu lassen.

4. **Versuchsergebnisse.** — a. *Ermittlung der Gesetzmässigkeiten für Lichtblitze in Schwellenwertnähe.* Es wurde mit einer gleichbleibenden Leuchtdichte der Mattscheibe von 370 asb gearbeitet, bei einer Apertur der Versuchsblende von 3 mm ⌀. Diese Werte ergeben für die mittlere Beobachter-Entfernung von 9,25 m eine Beleuchtungsstärke im Auge der Beobachter von rund 1.10^{-5} Lux. Um in Schwellenwertnähe zu gelangen, mussten durch passende Wahl der Grauglasfilter Schwächungen der Lichtblitze um die Faktoren 10^{-1} bis 10^{-3} vorgenommen werden.

Zu jeder Zeiteinstellung am Verschluss wurden daher Filterkombinationen benutzt, die einen Bereich von etwa 2 Grössenordnungen für die Lichtschwächung umfassten, so dass die Schwellenintensität mit Sicherheit erreicht werden konnte. Dabei kamen für jede Filterkombination 5-12 Lichtblitze gleicher Dauer und Intensität zur Auslösung, vermehrt um eine Anzahl blinder Blitze, die wahllos eingestreut wurden.

Jeder Beobachter notierte seine Ergebnisse im Dunkeln nach einer verabredeten Methodik. Damit liess sich eine gegenseitige Beeinflussung der Beobachter vermeiden ; es bestand auch keine Möglichkeit, sich etwa nach anderen Gesichtspunkten hinsichtlich der mutmasslichen Beobachtbarkeit unbewusst zu orientieren.

Für jede einzelne Einstellungsart — also z. B. Verschlusszeit 1/50 sec und Filterkombination 3 + 4 (d. h. für Durchlässigkeitsmoduln

$$\theta_3 \cdot \theta_4 = 1{,}63.10^{-1} \times 2{,}76.10^{-1} = 4{,}5 \cdot 10^{-2})$$

— wird aus den notierten wahrgenommenen Lichtblitzen aller Beobachter der Mittelwert gebildet und daraus die Beobachtbarkeit dieser Lichtblitze in Prozenten der echten Blitzzahl (echte Blitzzahl = 100 %) bestimmt.

In Abb. 2 sind die aus mehreren solchen Untersuchungsreihen berechneten Mittelwerte eingetragen, wobei von den Werten um 100 % und um 0 % der grösste Teil weggelassen werden konnte. Die Kurven in Abb. 2 stellen das Endergebnis aus etwa 5.000 einzelnen Beobachtungsvorgängen dar. Sie zeigen einen nahezu gleichförmigen Verlauf. Die Stellen, an denen sie die gestrichelt eingezeichnete Linie für eine 50 % — ige Beobachtbarkeit schneiden, sind besonders gekennzeichnet. Sie stellen, entsprechend unserer Definition nach Abs. 2 g, den gesuchten Schwellenwert für jede Verschlusseinstellung dar.

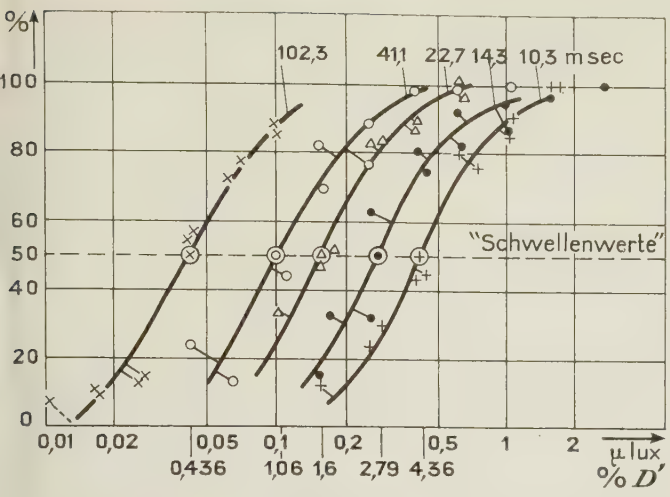

ABB. 2. — Die « Schwellenwerte » als Funktion der Beleuchtungsstärke bei verschiedenen Belichtungszeiten für « weisses » Licht.

Mit Rücksicht auf die Lamellenlaufzeiten des Compur-Verschlusses, die bei einigen msec liegen, wurde die niedrigste Blitzdauer mit ungefähr 10 msec gewählt. Die Eichung des Verschlusses geschah mittels Photozelle und Kathodenstrahl-Oszillographen.

Die Auswertung der Oszillogramme lieferte nach Mittelwertsbildung über jeweils 10 Aufnahmen mit der gleichen Belichtungsdauer die für die benutzten 5 Zeiteinstellungen des Verschlusses Abbildung 2 in msec als Parameter eingetragenen Blitzzeiten.

Zur Beurteilung farbiger Lichtsignale kann man sich aufgrund energiemässiger. Überlegungen nicht mehr der üblichen, für weisses Licht eingeführten Masseinheit der Beleuchtungstärke, des Lux, bedienen. Man müsste vielmehr unter Anwendung der spektralrelativen Werte des mechanischen Lichtäquivalents auf das Watt als Masseinheit für die Schwellenwerte übergehen. Dies aber würde zu wenig anschaulichen Ergebnissen führen. Andererseits handelt es sich hier darum, die Helligkeiten bezw. die Sichtbarkeit verschiedenfarbiger Signallichtquellen miteinander zu vergleichen. Dazu bedarf es eines Masstabes, der die durchschnittliche spektralrelative Augenempfindlichkeit berücksichtigt und dadurch in seiner Anwendung zu Ergebnissen führt, die vom Farbensinn des einzelnen, normal farbentüchtigen Beobachters unabhängig sind. Der Weg hierzu führt über die spektrale Energieverteilung der Lichtquelle und die spektralrelative Durchlässigkeit der benutzten Signalgläser.

Lassen wir bei unseren Lichtblitzgebern zunächst die Mattscheibe unberücksichtigt, so ist die « Helligkeit » H des auf das vorgesetzte Farbfilter fallenden Lichtes durch die Energie E der Lichtquelle und die Augenempfindlichkeit ε bestimmt (sichtbarer Bereich), also durch den Ausdruck :

$$H \int_{0,4}^{0,7} E_\lambda \, \varepsilon_\lambda \, d\lambda.$$

Nach Durchgang durch das Signalglasfilter wird dieser Ausdruck zu

$$H' = \int_{0,4}^{0,7} E_\lambda \, \varepsilon_\lambda \, D_\lambda \, d\lambda,$$

wenn D_λ die Durchlässigkeit des Filters für die einzelnen Wellenlängen bedeutet. Wählen wir nun als Vergleichsmasstab das Verhältnis $H'/H = D$, so ergibt sich damit die Gesamtdurchlässigkeit des Farbfilters ; und dieser Wert ist nunmehr unabhängig vom Farbensinn des Beobachters. Man kann die für verschiedenfarbige Filter erhaltenen D-Werte unmittelbar miteinander vergleichen und hat mit jedem einzelnen D-Wert sofort das Verhältnis zum normalen, weissen Bezugslicht, d. h. zu der gleichen Lichtquelle ohne vorgesetzte Farbenglasfilter. Für eine Bestimmung der Absolutwerte von H und H' wäre noch der Schwächungsfaktor der Mattscheibe in Rechnung zu setzen.

Es wurden die folgenden Farbenglasfilter untersucht :

(1) Signalrot, Filterdicke = 2,17 mm,
(2) Signalgelb, Filterdicke = 2,18 mm,
(3) Signalgrün, Filterdicke = 2,07 mm.

Dabei handelt es sich um die für Signalzwecke im deutschen Verkehrswesen amtlich zugelassenen Glassorten, hergestellt von der *Deutschen Spiegelglas-AG.*, Werk Mitterteich. Für die Beleuchtung der Mattscheibe wurde eine Wolframwendellampe in Sonderausführung nach Soang (Type H 15) von der Fa. Osram GmbH. verwendet. Die Betriebsspannungen betrugen 4 bezw. 5 Volt (± 3% Fehler), entsprechend einer wahren Farbtemperatur von 2180°K bezw. 2300°K (Angaben der Fa. Osram).

Die Art und Weise der Beobachtungen geschah in analoger Form zu denjenigen bei weissem Licht. Die Farbenglasfilter wurden jeweils in den am Verschluss anschraubbaren Tubus (R in Abb. 1) eingesetzt.

Das Ergebnis der Beobachtungen ist in einem Beispiel in Abb. 3 für Signalgrün dargestellt. Für Signalgelb und -rot ergaben sich ähnlich verlaufende Kurvenscharen, die jedoch nach höheren D' -Werten hin verschoben sind. Diese D' -Werte stellen die in Prozenten ausgedrückten Durchlässigkeitszahlen für die NG 5 -Grauglasfilter dar. Der Wert 100% bedeutet also, dass kein NG 5 -Schwächungsfilter mehr vorgesetzt ist, die Lichtdurchlässigkeit also 100%, die -schwächung 0% beträgt.

Die Kurven der Abbildungen 2 u. 3 geben also an, bei welchen Durchlässigkeitszahlen die Schwellen der Wahrnehmbarkeit für die verschiedenen Belichtungszeiten erreicht wurden. Während bei Signalgelb und -rot im Durchschnitt wiederum, wie bei weissen Lichtblitzen, 4-5000 Beobachtungsvorgänge zur Auswertung genügten, mussten bei Signalgrün über 9000 einzelne Beobachtungswerte gesammelt werden.

Wir wollen nun prüfen, ob und inwieweit das ursprünglich für eine rein foveale Beobachtungsweise ermittelte BLONDEL u. REY'sche Gesetz auch in dem hier untersuchten Falle gültig ist. Zu diesem

Zweck wurden diejenigen einander zugeordneten Werte von B und t unserer Messungen ausgesucht, die für die Schwellenwahrnehmbarkeit, d. h. für eine 50 % -ige Beobachtbarkeit, gefunden worden waren.

Wählen wir t als Parameter, so können die zugehörigen Werte für B den für die verschiedenen Signalfarben ermittelten Kurvenscharen (z. B. den Abb. 2 u. 3) entnommen werden ; es sind dies die für jedes t durch die Schnittpunkte der 50 % -Linie der Beobachtbarkeit bestimmten Abszissenwerte in µlux bezw., wegen der Umrechnungsschwierigkeiten bei farbigem Licht, in dem der Beleuchtungsstärke B äquivalenten D'-Masstab. Diese 5 Wertepaare für B und t sind in Abbildung 4 für die 4 untersuchten Signalfarben eingezeichnet. Ausserdem konnten noch Wertepaare mit B bezw. D' als Parameter ermittelt werden. Zu diesem Zweck wurden alle Werte für die Beobachtbarkeit (in %) aus den Beobachtungsprotokollen zusammengestellt, die jeweils für die gleiche Beleuchtungsstärke B (D') gelten. Durch Interpolation kann man hieraus für die einzelnen B- (D'-) Werte die zugehörigen t-Werte für eine 50 % -ige Beobachtbarkeit ermitteln.

Die so farbige Signal-Lichttblitze erhaltenen Wertepaare für D' und t für sind ebenso wie diejenigen, welche für weisses Licht (B und t) bestimmt wurden, in der Abbildung 4 eingetragen. Die für weiss in dieser Zeichnung eingetragenen Kurven gelten sowohl für den rechten Masstab (B in µlux) als auch für den linken (D' in %). Wie man sieht, kann man jeweils durch die zusammengehörigen Punkte eine unter 45° geneigte Gerade ziehen. Da für beide Koordinaten ein logarithmischer Masstab gewählt wurde, bedeuten diese 45° -Geraden somit bereits die Erfüllung des Gesetzes nach dem vereinfachten Ausdruck $B.t$ = const. → $D'.t$ = const.

Es lässt sich nun zeigen, dass der Streubereich der eingetragenen Werte das Blondel und Rey'sche Gesetz auch in seiner allgemeinen Form

$$B \,.\, t = B_0(t + \alpha)$$

umfasst. Dafür wurde als Mittelwert α = 210 msec von den Autoren Blondel und Rey übernommen. B_0 ist die Beleuchtungsstärke an der Schwelle der Wahrnehmbarkeit für Dauerlicht.

Setzen wir für weiss an : B_0 = 0,0175 µlux*, so fällt die Kurve für

$$B \,.\, t = B_0(t + \alpha) = 0{,}0175\ (t + 210)$$

etwa in der Mitte des zulässigen Belichtungszeiten-Intervalls mit der Geraden

$$B.t \sim D'.t = \text{const.}$$

zusammen (siehe Abb. 4).

Auf einen analogen Ansatz für die Signalfarben kann verzichtet werden. Es würde sich der gleiche Kurvenverlauf wie bei weiss ergeben, nur mit dem Unterschied, dass die Kurven in der Zeichnung nach

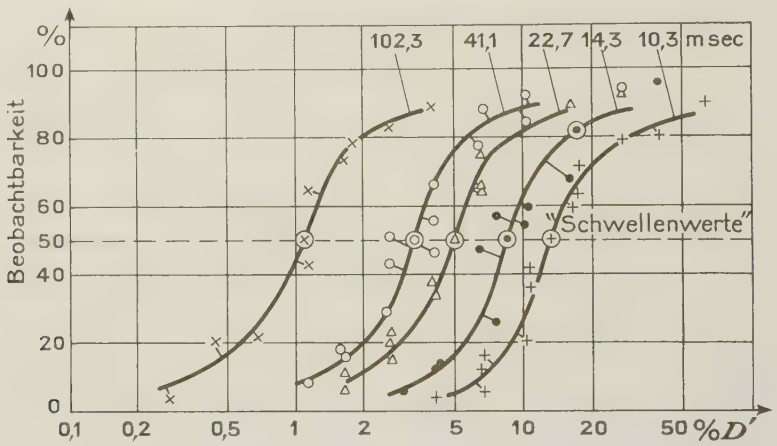

Abb. 3. — Die « Schwellenwerte » als Funktion der Beleuchtungsstärke bei verschiedenen Belichtungszeiten für Signalrgün (Glasdicke = 2,07 mm).

oben hin parallel verschoben wären. Wir können uns also auf die bereits für weisse Lichtblitze hergeleitete Kurve beziehen. Wie der Vergleich zeigt, fällt auch hier das erweiterte Gesetz in den Streubereich der ermittelten Werte. Somit ergibt sich das folgende Resultat :

Die Sichtbarkeit kurzdauernder ($t \leqslant 0{,}1$ sec), an der Schwelle der Wahrnehmbarkeit liegender Lichtsignale folgt auch dann dem Gesetz von Blondel und Rey

$$B \,.\, t = B_0(t + \alpha),$$

wenn *a*) nicht streng foveal beobachtet wird,
b) der Schwellenwert durch eine sog. 50 %-ige Beobachtbarkeit definiert wird,

* Wegen der nicht rein fovealen Beobachtungsweise war eine experimentelle Prüfung dieses B_0-Wertes nicht möglich. Er stimmt aber z. B. gut mit Angaben von Borchardt [40] überein, die für eine foveale Beobachtung ohne Fixationspunkt nach längerer Dunkeladaption gelten. Aus seinen Daten errechnet man einen Schwellenwert von B_0 = 0,011 µ lux. Das Netzhautbild war mit 0,4 mm Durchmessen etwas grösser als die fovea centralis (ca. 0,2-0,3 mm ∅), woraus sich die gegenüber dem rein fovealen Schwellenwert [41] erhöhte Empfindlichkeit erklärt. In unserem Falle trägt der extrafoveale Beobachtungsanteil ebenfalls zu einer gegenüber dem fovealen Sehen gesteigerten Empfindlichkeit bei.

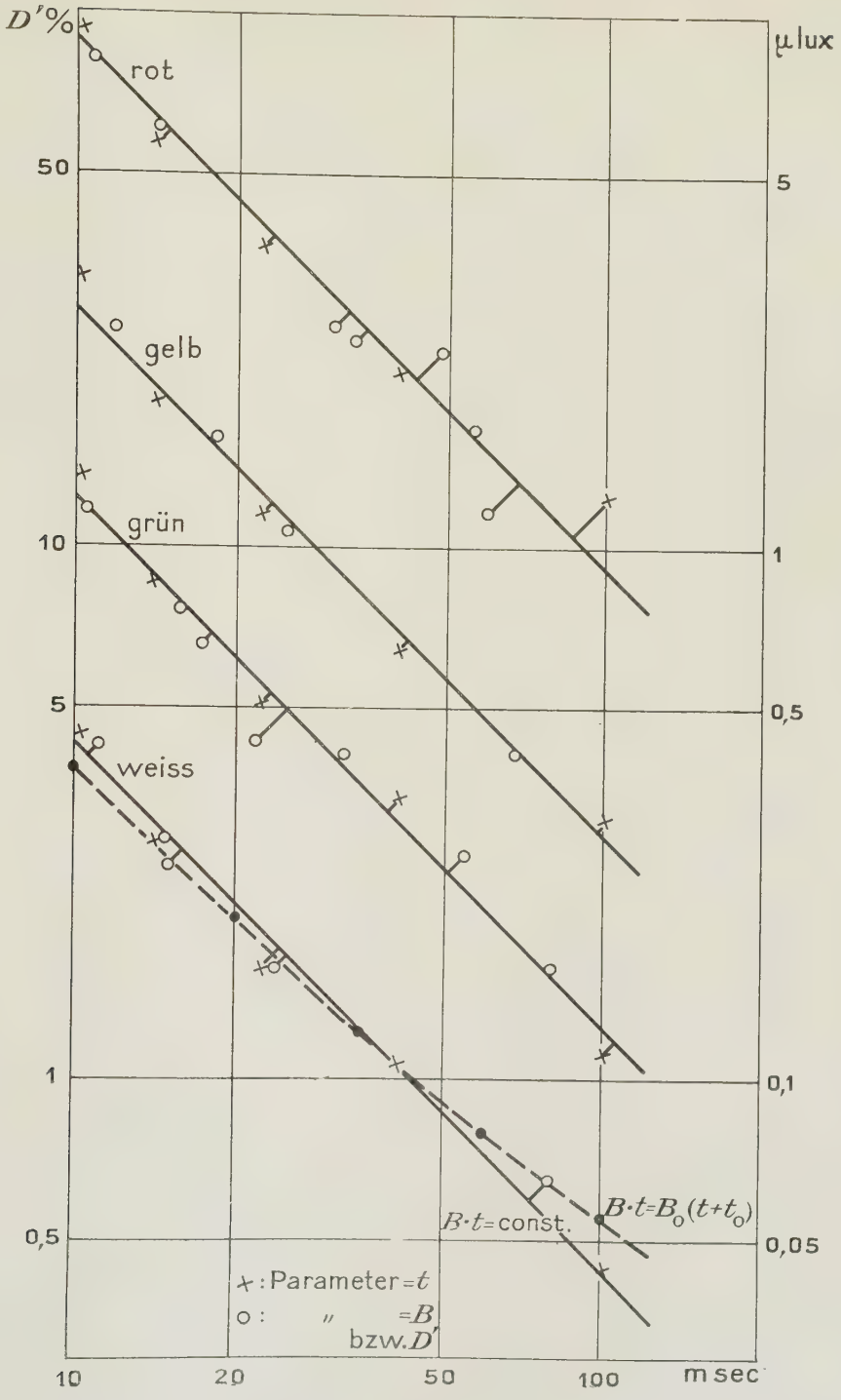

Abb. 4. — Gesetz von Blondel und Rey ; Prüfung für « weisse » und farbige Lichtblitze.

c) die im Verkehrswesen üblichen Signalfarben benutzt werden. Es ist zulässig, hierauf auch die vereinfachte Form dieses Gesetzes : $B \, . \, t = \text{const.}$ anzuwenden.

b. *Einfache und spezifische Schwellenwerte.* — Die Gültigkeit des Blondel und Rey' schen Gesetzes in der einfachen Form $B.t = \text{const.} \rightarrow D' \, . t = \text{const.}$ erlaubt es, die Schwellenwerte direkt aus den Geraden der Abbildung 4 zu bestimmen. Für weisses Signallicht erhalten wir somit unter den in Abs. 2 a-g festgelegten Versuchsbedingungen einen einfachen Schwellenwert von

$$S = B \, . \, t = 4{,}3 \, . \, 10^{-9} \text{ luxsec.}$$

Zur Bestimmung der spezifischen Schwellenwerte bedienen wir uns des für die farbigen Lichtsignale eingeführten Vergleichsmasstabes D (s. Abs. 4 a). Die Brauchbarkeit dieser D-Werte liegt darin, dass sie zu Aussagen darüber herangezogen werden können, um welchen Faktor die Helligkeit H' für ein bestimmtes Farbfilter verändert werden muss, um eben für dieses Filter den Schwellenwert zu erreichen. Man braucht dazu nur den reziproken Wert $D^+ = 1/D$ zu bilden. D^+ gibt also den Faktor an, um welchen die Strahlungsenergie E des Lichtstromes der Lichtquelle selbst oder, wenn mit einer Mattscheibe gearbeitet wird, die Leuchtdichte der Mattscheibe erhöht werden muss, um bei sonst unveränderten Bedingungen eine gegenüber dem weissen Licht der Lichtquelle etwa gleich gute Sichtbarkeit im Bereich ein farbiges Signalglasfilter bestimmter Dicke gesetzt wird. Dieser Faktor ist aber genau der gleiche, mit dem man in der Gleichung für D (s. Abs. 4a) H' multiplizieren muss, um den Wert für H zu erhalten (bei der gleichen Glasdicke), d. h. um mit und ohne Farbglasfilter unabhängig vom Farbensinn des Beobachters gleiche Helligkeit zu erhalten, also :

$$H = \frac{1}{D} H' \text{ mit } \frac{1}{D} = \frac{D'(\Delta\lambda)}{D'(\mathrm{w})} = D^+$$

für jede bestimmte Lichtblitz-Dauer.

Die auf diese Weise aus der Abb. 4 ermittelten D^+-Faktoren sind der Tabelle I aus der obersten Zeile unter der Spalte « experimentell » zu entnehmen. Berücksichtigt man, dass die Glasdicke für alle 3 Signal-

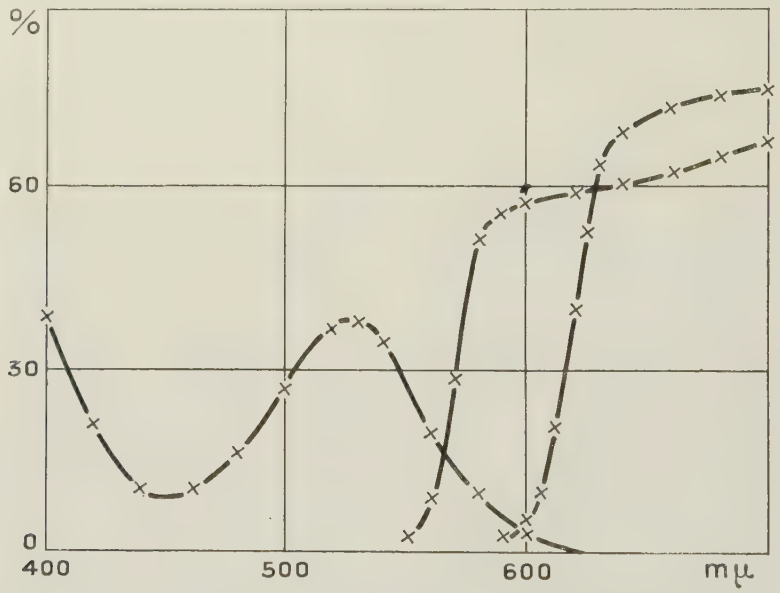

Abb. 5. — Spektrale Durchlässigkeit der Signalglas-Filter.

der Schwellenwerte (also bei sehr schwachen und kurzen Lichtblitzen) zu erhalten. Bei der Ausführung der Beobachtungen geschah dies in der Weise, dass die für weisses Licht vorgesetzten NG5-Grauglasfilter durch schwächere Filter mit entsprechend grösserer Durchlässigkeit ersetzt wurden, und zwar bei den gleichen Zeiteinstellungen am Compurverschluss, wie für die Untersuchung ohne Farbfilter.

Die Ermittlung der D^+-Faktoren geschieht nun durch die Interpolation aus den Geraden in Abbildung 4 unter Ansatz der D'-Werte. Diese D'-Werte geben den prozentualen Anteil des jeweils von der NG5-Filterkombination durchgelassenen Lichtstromes der Lichtquelle an. Bildet man — im Bereich $\Delta\lambda$|—|das Verhältnis der D'-Werte zwischen farbigem Signallicht $D'(\Delta\lambda)$ und weissem Licht $D'(\mathrm{w})$, dann erhält man damit den Faktor, um den man den Lichtstrom erhöhen muss, damit wieder Schwellenintensität herrscht, wenn vor das weisse Signallicht farben nahezu konstant war, so fällt besonders auf, dass gerade die Farbe Signalgrün den niedrigsten Schwellenwert hat.

In der Abbildung 5 sind die Kurven für die spektrale Durchlässigkeit der 3 Signalfarbenfilter eingezeichnet, und zwar unter Berücksichtigung der Glasdicke der benutzten Farbfilter und ihrer Reflexionskoeffizienten R_d. Während die Farbe Signalgrün — vom violetten Gebiet abgesehen — ein Maximum der Durchlässigkeit bei einer Wellenlänge von etwa 525 mμ besitzt, ist die Farbe Signalgelb durch eine steile Absorptionskante bei etwa 570 mμ gekennzeichnet. Man würde also den niedrigsten Schwellenwert eher für Signalgelb als für Signalgrün erwarten, weil ja die grösste Augenempfindlichkeit etwa bei 555 mμ liegt.

In der Abbildung 6 sind die Kurven für die Augenempfindlichkeit eingetragen. Die ausgezogene Kurve stellt die international festgelegte Augenempfindlich-

keit für das Tagessehen, also für die Zäpfchen — und damit für ein streng foveales Sehen — dar. Die gestrichelte Kurve ist diejenige für das Dämmerungssehen (nach KOHLRAUSCH [33-35]).

Ausserdem wurden in Abb. 6 die Kurven für die Strahlungsenergie der Wolframwendellampe H 15 im sichtbaren Gebiet bei einer wahren Temperatur von 2180°K und von 2300°K, entsprechend den benutzten Betriebsspannungen von 4 bezw. 5 Volt eingezeichnet. Ihrer Berechnung lag das Plancksche Strahlungsgesetz zugrunde, dessen einzelne Werte E_λ für die verwendete Sonderlampe mit dem Faktor 0,45 zu multiplizieren sind (lt. Angabe des Herstellers), um die zunächst für schwarze Temperaturen geltenden Kurven auf solche für die wahren Farbtemperaturen der Wolframwendellampe zu transformieren. Die Berücksichtigung dieser Energieverteilung der Lichtquelle selbst lässt erst recht die beste Beobachtbarkeit für Signalgelb erwarten — jedenfalls solange die Tageskurve der Augenempfindlichkeit hier als gültig angesehen wird.

Nun sind die Verhältnisse in der Praxis so gelagert, dass wohl nur in Ausnahmefällen die Lage des Signalgebers exakt bekannt ist. Vielmehr dürfte meistens nur die ungefähre Richtung festliegen, in der man die Lichtquelle zu erwarten hat, z. B. die Lage einer Verkehrsampel, die Leuchtfeuer von Flugplätzen usw. Dann aber wird man auch nur teilweise foveal beobachten können, jedenfalls solange der Standort des Signallichtes noch genau ausgemacht werden muss. Darauf aber kommt es ja bei der Sichtbarkeit von Signallichtquellen wesentlich mit an. Gemäss Abs. 2 d hatten wir daher unsere Versuchsbedingungen diesen Verhältnissen angepasst. Es erscheint somit zweifelhaft, ob es berechtigt ist, für die Beurteilung der Sichtbarkeit von Signallichtquellen unter den angenommenen allgemeinen Beobachtungsverhältnissen die international festgelegte spektralrelative Augenempfindlichkeit als Vergleichsgrundlage zu wählen.

Zur Prüfung dieser Frage wurde daher das Problem theoretisch durchgerechnet, und zwar sowohl für das reine Zäpfchen-Sehen als auch für das Dämmerungs-Sehen. Es wurden also in unserem Falle die Farbfilter-Kurven der Abbildung 5 und die Kurven für die Zäpfchenempfindlichkeit sowie die Strahlungsverteilung der Lichtquelle bei 2180 °K aus Abbildung 6 miteinander kombiniert und daraus die 4 resultierenden Kurven — für weisses, rotes, gelbes und grünes Lichtermittelt (siehe Abb. 7). In gleicher Weise wurden diese Kurven nunmehr auch für die Dämmerungswerte aus Abb. 6 bestimmt (Ergebnis s. Abb. 8). Ausserdem wurden dieselben Operationen mit der Tageskurve nochmals für die Strahlungsverteilung der Lichtquelle bei 2300°K (aus Abb. 6) ausgeführt, um festzustellen, ob sich in den Relationen ein wesentlicher Unterschied ergibt, wenn man die Betriebsspannung der Lichtquelle von 4 auf 5 Volt erhöht. Diese Spannungserhöhung hatte sich bei den Untersuchungen mit Signalrot als not-

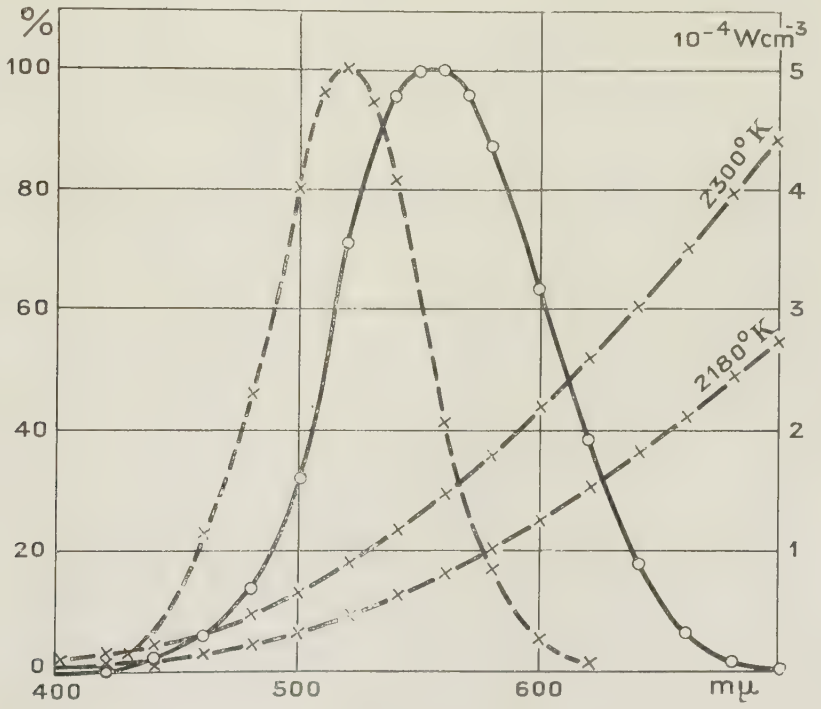

ABB. 6. — Kurven für die Augenempfindlichkeit bei Hellund Dunkeladaption und Energieverteilung der Wolframwendellampe bei verschiedenen Farbtemperaturen.

wendig erwiesen, da die Leuchtdichte der Mattscheibe bei 4 Volt für die Kurzen roten Lichtblitze bereits zu gering war*.

Zur Bestimmung der D^+-Faktoren wurden die Integrale für H und H' auf graphischem Wege durch Planimetrierung der Kurven aus den Abbildungen 7 u. 8 bestimmt. Das Ergebnis ist mit den bereits gewonnenen experimentellen Resultaten in der Tabelle I in der obersten Zeile zusammengestellt.

Bereits der Vergleich D^+-Faktoren, die untereinander nicht übereinstimmen, beweist die Berechtigung der Annahme, dass der Ansatz der Tageskurve für die Augenempfindlichkeit in diesem Falle unzulässig ist. Ebensowenig aber kann man an ihrer Stelle die Dämmerungskurve verwenden, weil ja keineswegs alle Lichteindrücke von den Beobachtern extrafoveal aufgenommen wurden. Auch dies findet man in der obersten Zeile der Tabelle I bestätigt.

Die experimentell ermittelten D^+- Faktoren resultieren also aus einem gemischten Zäpfchen- und Stäbchensehen.

Der Vergleich der experimentell bestimmten D^+-Faktoren mit den berechneten bleibt jedoch insofern unbefriedigend, als je nach der Signalfarbe einmal die Art des Tagessehens und einmal die des Dämmerungssehens vorzuherrschen scheint. Da jedoch alle Beobachtungen unter genau den gleichen Bedingungen stattfanden, ist dies gänzlich unwahrscheinlich.

Um diese Frage zu klären, möge zunächst auf einige bereits bekannte Erkenntnisse hingewiesen werden. Beim Dämmerungssehen ist die Lichtempfindlichkeit der Retina in der Netzhautgrube am geringsten. Schon bei einem Abstand von 1,2° davon steigt sie jedoch auf das Doppelte ; bei 4° Abstand hat sie bereits den 60fachen Betrag erreicht! Ferner beruht das Dämmerungssehen auf der Aktivität der Stäbchen, die jedoch keine Farben, sondern nur Helligkeitswerte zu unterscheiden vermögen. Die Dämmerungskurve des Auges (s. Abb. 6) gilt also nicht etwa für eine Farbenerkennung, sondern nur für die Lichtreizung der Retina in Abhängigkeit von der Wellenlänge. Eine spezifische Farbenbewertung der die Reizung auslösenden Strahlung ist mittels der Stäbchen-nicht möglich. Sobald die Reizung der Retina jedoch so stark wird, dass Farben erkannt werden, ergibt sich bereits eine Translation der Dämmerungskurve in Richtung auf diejenige für das Zäpfchensehen.

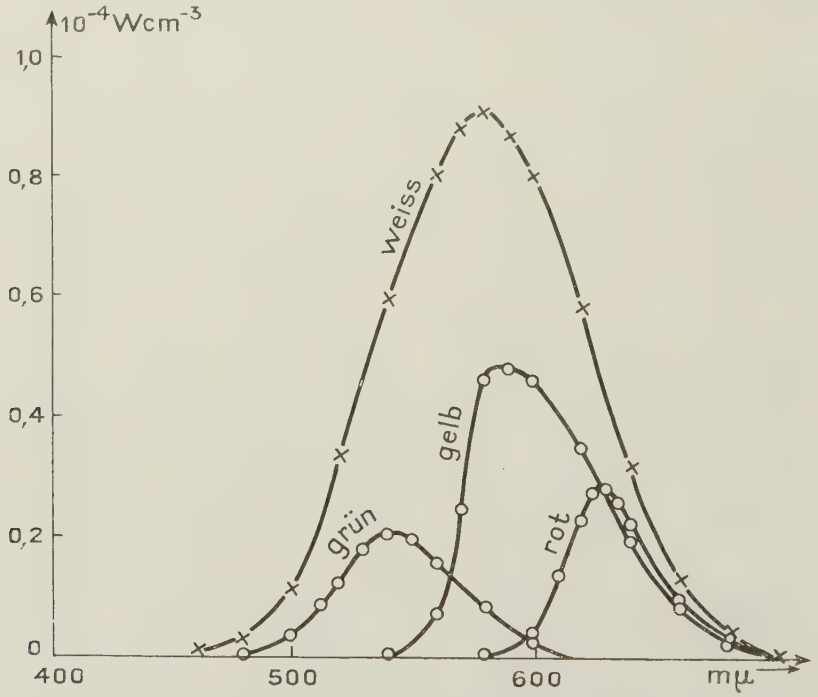

Abb. 7. — Die Augenempfindlichkeit bei Helladaption bezogen auf das Licht einer Wolframwendellampe (2180°K) mit und ohne Signalglasfilter.

Bildet man nun aus den D^+-Werten der Tabelle I die Verhältnisse zwischen den experimentell bestimmten und den berechneten (2. Zeile der Tabelle), so erkennt man, dass sich für die Tageskurve bei Signalgelb mit 2,4 und bei Signalrot mit 3,2 nahezu gleich grosse Verhältniszahlen ergeben, während die entsprechenden Relationen bei der Dämmerungskurve weit auseinanderliegen. Die experimentellen Resultate weisen also für Signalgelb und -rot ein charakte-

* Es ergab sich, dass eine Berücksichtigung der erhöhten wahren Farbtemperatur auf 2300 °K bei den graphischen Auswertungen für Signalrot nicht erforderlich war.

TABELLE I

Vergleich d. experimentell ermittelten u. d. berechneten D^+-Faktoren

Faktoren	experimentell			ber., Tageskurve			ber., Dämmerungskurve		
	2180°K		2300°K	2180°K		2300°K	2180°K		
	grün	gelb	rot	grün	gelb	rot	grün	gelb	rot
D^+	3,0	6,8	22	7,13	2,85	6,51	(3,57)	13,36	237
$\frac{D^+_{exp.}}{D^+_{ber.}}$	—	—	—	0,4	2,4	3,2	—	0,5	0,09
$\frac{D^+_{rot}}{D^+_{gelb}}$	3,2			2,3			16,5		

ristisches Merkmal des Tagessehens auf, was darauf hindeutet, dass die Beobachtungen vorwiegend foveal stattfanden. Noch deutlicher erkennt man das, wenn man das Verhältnis der D^+-Faktoren für Signalrot zu Signalgelb bildet. Diese in der letzten Zeile der Tabelle I eingetragenen Verhältniszahlen zeigen die relativ gute Übereinstimmung der für das Tagessehen berechneten mit den beobachteten Werten, während der Verhältniswert für die Dämmerungskurve völlig herausfällt.

Dass die experimentell gewonnenen D^+-Faktoren für die Farben rot und gelb grösser sind als die nach der Tageskurve berechneten, ergibt sich aus folgender Überlegung nun ohne weiteres :

Extrafoveal « beobachtete » rote oder gelbe Lichtblitze wurden — soweit sie sehr lichtschwach waren, also in Schwellenwertnähe lagen — meistens garnicht erkannt! Die Stäbchen sprechen nämlich auf die Reizungen aus diesem Spektralbereich bei rot praktisch überhaupt nicht, bei gelb nur unbedeutend an (vgl. Abb. 5 u. 6). Die Zäpfchenempfindlichkeit aber ist in der fovea centralis am grössten und nimmt zur Peripherie der Retina hin immer mehr ab. Jeder extrafoveal nicht erkannte Lichtblitz (der aber foveal erkannt worden wäre!) dieser beiden Signalfarben bedingt eine geringere Sichtbarkeit, also eine nahezu konstante Verschiebung der D^+-Faktoren nach höheren Werten, d. h. nach grösseren Durchlässigkeiten der NG5-Filter ; das kommt aus der Tab. I auch sehr gut zum Ausdruck.

Es bleibt noch zu klären, warum sich gerade bei Signalgrün ein experimentelles Ergebnis einstellt, welches auf reines Dämmerungssehen hinzudeuten scheint. Zunächst erkennt man aus dem Vergleich der Kurve für Signalgrün (Abb. 5) mit den beiden Augenkurven (Abb. 6), dass jetzt allerdings jeder extrafoveal beobachtete Lichtblitz eine hervorragende Aufnahmefähigkeit der Retina antrifft. Es fällt nämlich das Maximum der Augenempfindlichkeit für die Stäbchen und die maximale Filterdurchlässigkeit für Signalgrün fast zusammen! Jeder extrafoveal auf die Retina einfallende grüne Lichtstrahl wird wesentlich besser erkannt, als bei fovealer Beobachtung. Die Empfindlichkeit der Stäbchen steigert sich dabei mit wachsendem Einfallswinkel erheblich. Es ist somit verständlich, dass bei Signalgrün jeder extrafoveal beobachtete Lichtblitz das Ergebnis dahingehend beeinflusst, dass sich niedrigere D^+-Faktoren ergeben müssen, als bei streng fovealem Sehen erwartet werden kann. Der Einfluss des Dämmerungssehens ist also evident.

Warum aber ist der Faktor derart niedrig, nämlich 3,0, dass anscheinend überhaupt nicht foveal beobachtet wurde? Die Antwort darauf erhält man, wenn man die von PIÉRON (37-39) gefundenen Resultate hier anwendet. Er stellte nämlich fest, dass die

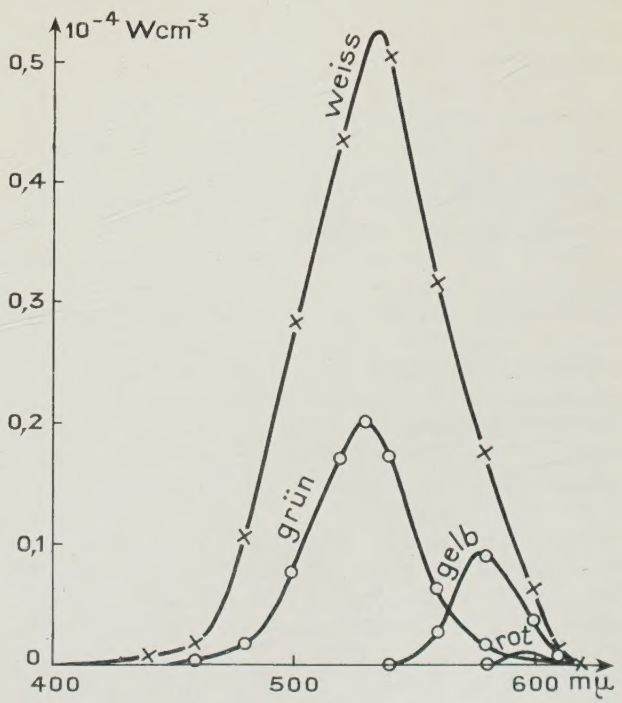

ABB. 8. — Die Augenempfindlichkeit bei Dunkeladaption bezogen auf das Licht einer Wolframwendellampe (2180°K) mit und ohne Signalglasfilter.

Empfindlichkeit der Stäbchen bei kurzzeitigen Lichterscheinungen wesentlich schneller ansteigt als diejenige der Zäpfchen. So ist das Verhältnis z. B. bei einer Lichtblitzdauer von 800 msec bereits 90 : 1, bei 5 msec sogar 315 : 1! Die bei unserer Berechnung benutzten Kurven der Abb. 6 gelten jedoch für Dauerlicht. Der Einfluss des extrafovealen Beobachtungsanteils für das grüne Signallicht erhöht sich damit noch wesentlich, da die benutzten Beobachtungszeiten zwischen 10 und 100 msec lagen.

Weiter haben BROCA und SULZER (28-30) nachgewiesen, dass der Anstieg der Lichtempfindlichkeit für blaues Licht ganz besonders steil verläuft, wobei das Maximum etwa im Bereich von 40-50 msec erreicht wird. Auch sie fanden für die Stäbchen eine wesentlich geringere Einstellzeit auf volle Empfindlichkeit als für die Zäpfchen. Die Dämmerungskurve würde somit je nach der Dauer der Lichtblitze einen etwas anders verlaufenden, sehr steilen Anstieg haben, der schon bei kurzen Wellenlängen, also etwa bei 400-500 mμ wesentliche Werte erreicht.

Eine derart für kurze Lichtblitze korrigierte Dämmerungskurve würde also im extrafovealen Beobachten lediglich für das Signalgrün einen Gewinn bringen, hier allerdings einen beträchtlichen. Denn im kurzwelligen Spektralintervall, zwischen 400 und 500 mμ, kommt die Veränderung der Dämmerungskurve spektralrelativ besonders stark zur Geltung. Hier ist nämlich der Anteil der Energieverteilung der Lichtquelle noch nicht erheblich — er liegt hauptsächlich im Rot — während der Zuwachs sich nicht nur auf das Maximum des grünen Signalfilters bei ca. 525 mμ sondern nunmehr noch zusätzlich auch auf dessen beträchtliche Durchlässigkeit für blau voll auswirkt.

Der D^+-Faktor der Tabelle I für grün nach der Dämmerungskurve (3,57) ist also viel zu hoch ausgefallen. Eine genaue Berechnung ist zwar hier nicht möglich, jedoch dürfte der vorstehend qualitativ erbrachte Nachweis, dass der Faktor wesentlich geringer sein muss, genügen.

Damit aber ist geklärt, dass der experimentell bestimmte D^+-Faktor für grün keineswegs auf eine überwiegend extrafoveale Beobachtungsweise hindeutet, womit der scheinbare Widerspruch zu den bereits für Signalgelb und-rot gewonnenen Erkenntnissen beseitigt ist. Dieses Resultat steht zudem mit Kontrollversuchen im Einklang, die unter Beibehaltung der Versuchsbedingungen einen durchschnittlichen extrafovealen Beobachtungsanteil von etwa 25-30 % ergaben.

Zur Prüfung der Frage, ob im fovealen Sehen ein farbloses Intervall vorhanden ist oder nicht, war ein zweiter Lichtblitzgeber der gleichen Konstruktion, wie in Abb. 1 wiedergegeben, unmittelbar neben dem eigentlichen Versuchs-Blitzgeber aufgebaut worden. Dieser zweite Blitzgeber diente dazu, den Beobachtern exakte Vergleichsfarben zu liefern. Eine solche Einrichtung erwies sich als nützlich, um bei den grossenteils sehr schwachen Lichtblitzen ein möglichst genaues Urteil über den beobachteten Farbton der Blitze zu erhalten. Es wurde dabei so verfahren, dass bei jeder einzelnen Einstellung — sowohl für die Versuchsdauer als auch für die Filterkombination — unmittelbar vor Beginn und nach Beendigung einer Gruppe von Lichtblitzen ein genügend langer Vergleichsblitz gegeben wurde.

Nach den bereits zitierten Ergebnissen von KÖNIG (22) sollen für den überwiegend fovealen Anteil der Beobachtungen alle Farben ausser gelb beim Erreichen des Schwellenwertes sofort über die Schwelle treten. Hiermit stehen unsere Beobachtungen in befriedigendem Einklang. Der wetaus grösste Teil aller in der Farbe nicht richtig erkannten Lichtblitze wurde als weiss bis grau angesprochen, wobei diese Fehlurteile zu einem überwiegenden Anteil auf die älteren Personen unter den Beobachtern entfielen.

Bei Signalrot ergab sich eine ausgezeichnete Farberkennung. Die Fehlurteile blieben noch unter 10 %, und von 10 Beobachtern gab höchstens einer überwiegend falsche Farburteile ab. Einfacher und spezifischer Schwellenwert fallen für Signalrot also zusammen.

Signalgrün bereitete den Beobachtern die grössten Schwierigkeiten, was zweifellos auf die ausserordentlich gesteigerte Empfindlichkeit des extrafovealen Sehens und die damit verbundene Unsicherheit hinsichtlich der Farberkennung zurückzuführen ist. Trotzdem gaben nur etwa 20 % der Beobachter überwiegend falsche Farburteile ab. Es wurden auch nur etwa 20 % aller Lichtblitze nicht in der richtigen Farbe erkannt. Das Resultat spricht zwar für ein farbloses Intervall zwischen dem spezifischen und dem einfachen Schwellenwert beim extrafovealen Sehen, erscheint jedoch als zu unsicher, um daraus auch für die hier vorliegende kombinierte Beobachtungsweise ein farbloses Intervall herleiten zu können. Immerhin ergibt sich gegenüber Signalrot offensichtlich eine etwas geminderte Farben-Erkennbarkeit.

Etwas deutlicher tritt dieser Effekt bei Signalgelb auf. Hier wurden von ungefähr 30 % aller Beobachter überwiegend Farbfehlurteile abgegeben. Etwa 25-30 % aller Lichtblitze wurden als andersfarbig, davon etwa 80 % als farblos (weiss bis grau) angesprochen. Zweifellos ist also die Beurteilung der Farbe Signalgelb von allen drei Signalfarben in Schwellenwertnähe am unsichersten, was dem Ergebnis von KONIG nahekommt. Da der Schwellenwert jedoch jetzt für eine 50 %-ige Beobachtbarkeit definiert wurde, so reicht auch dieses Ergebnis kaum aus, um hierfür ein farbloses Intervall zwischen dem spezifischen und dem einfachen Schwellenwert als gesichert anzunehmen.

LITERATUR ÜBERSICHT

[1] BLONDEL & REY, *Journ. Phys. Radium*, **1**, 1911, 530, 643.
[2] BLONDEL & REY, *C. r. Ac. Sci.*, **153**, 1911, 54.
[3] BLONDEL & REY, *C. r. Ac. Sci.*, **162**, 1916, 587, 861.
[4] BLONDEL & REY, *C. r. Ac. Sci.*, **178**, 1924, 276.

[5] v. Kries, *ZS. f. Sinnesphysiol.*, **41**, 1907, 373.
[6] Gildemeister, *ZS. f. Sinnesphysiol.*, **48**, 1914, 252.
[7] Rutenburg, *ZS. f. Sinnesphysiol.*, **48**, 1914, 268.
[8] Blondel & Rey, *Journ. Phys.*, **1**, 1911, 643.
[9] Langmuir and Westendorp, *Physical.*, 1931, 273.
[10] Blondel, *C. r. Ac. Sci.*, **204**, 1937, 1796.
[11] Kurzke & Sattler, *Unveröffentlichte Arbeiten des Forsch.-Inst. f. Phys.*, 1942 und 1943.
[12] Piéron, *C. r. Ac. Sci.*, **170**, 1920, 525.
[13] Piéron, *C. r. Ac. Sci.*, **170**, 1920, 1203.
[14] Ricco, *Ann. di Ottal.*, **6**, 1877.
[15] Charpentier, *C. r. Ac. Sci.*, **91**, 1880, 995.
[16] Asher, *ZS. f. Biol.*, **35**, 1897, 394.
[17] Loeser, *Arch. f. Ophth.*, **60**, 1905, 97.
[18] Haupt, *Journ. of. exp. Psychol.*, **5**, 1922, 347.
[19] Abney, *Phil. Trans.*, **183**, 1892, 537.
[20] Abney, *Phil. Trans.*, **190**, 1897, 155.
[21] Abney, *Proc. Roy. Soc. London*, **87**, 1912, 147.
[22] König, *Berl. Ber.*, 1894, 577.
[23] Arndt, *ZS. f. techn. Phys.*, **8**, 1934, 300.
[24] v. Kries & Nagel, *ZS. f. Psych. Sinnesorgane* **23**, 1900, 161.
[25] Lummer, *Verh. d. D. Phys. Ges.*, **6**, 1904, 62.
[26] Rouse, *J. O. S. A.*, **42**, 1952, 626.
[27] Davy, *J. O. S. A.*, **42**, 1952, 937.
[28] Broca & Sulzer, *C. r. Ac. Sci.*, **134**, 1902, 831.
[29] Broca & Sulzer, *C. r. Ac. Sci.*, **136**, 1903, 1287.
[30] Broca & Sulzer, *C. r. Ac. Sci.*, **137**, 1903, 944, 977 1046.
[31] Piper, *ZS. f. Psych. u. Physiol. d. Sinnesorg.*, **31**, 1903, 161; **32**, 98, 161.
[32] Piper, *Klin. Monatsbl. f. Augenheilk.*, **45/I**, 1907, 357.
[33] Kohlrausch, *Pflügers Archiv.*, **200**, 1923, 210, 216.
[34] Kohlrausch, *Licht und Lampe* 1923, 555.
[35] Kohlrausch, *D. Opt. Wochenschr.*, **10**, 1924, 59.
[36] Hill, *Proc. Phys. Soc.*, **59**, 1947, 560, 574.
[37] Piéron, *C. r. Biol.*, **82**, 1919, 1162.
[38] Piéron, *C. r. Biol.*, **83**, 1920, 753 u. 1072.
[39] Piéron, *C. r. Ac. Sci.*, **180**, 1925, 462.
[40] Borchardt, *ZS. f. Sinnesphysiol.*, **48**, 1914, 176.
[41] v. Kries, *ZS. f. Physiol*, **15**, 1897.

Manuscrit reçu le 18 août 1957.

INFORMATIONS

A **Summer School** on :

I. *Optical image assessment using frequency response techniques* has been organized by Dr H. H. Hopkins with additional lectures given by Prof. Dr H. Frieser and Dr E. H. Linfoot, Monday, 7th July, to Thursday, 10th July, Imperial College of Science and Technology, London, SW 7.

A. *Syllabus* (7 th July - 10 th July).

(1) The use of frequency response methods in optics. (2) Calculation and measurement of response. (3) Aberration tolerances. (4) The use of the geometrical optical approximation. (5) Interferometric and scanning techniques for measurement of frequency response. (6) Relation to other methods of image-assessment. (7) Response techniques applied to photographic emulsions. (8) Information theory and its application to image-assessment.

II. *Elementary Fourier analysis* has also been organized Monday, 30th June, and Tuesday, 1st July.

B. *Fourier analysis* (30th June and 1st July).

An optional refresher course of six lectures : (1) Sine, cosine and exponential series. (2) Fourier integrals, the inversion theorem. (3) Fourier transforms; the shift, convolution and related theorems. (4) Functions of more than one variable. (5) Special cases : step and delta functions. (6) Relation to diffraction theory. Correlation functions.

The Summer School should be of interest to all those whose work is concerned with optical images. Enquiries should be addressed to Dr H. H. Hopkins at the above address.

The lectures will be given in the Physics Department, Imperial College, Imperial Institute Road, S.W. 7.

Application for admission should be made to the Registrar of the Imperial College, Prince Consort Road, S.W.7. The fee for the main course is £4. 40. ; for the refresher course, £1. 10. Students of the College and Inter-collegiate students will be admitted free to the lectures (on production of an inter-collegiate ticket).

Fifth meeting and conference of the International Commission for Optics (ICO) in Stockholm, 24th August - 30th August 1959.

Conference on " Modern systems for detecting and evaluating optical radiation ".

Meeting secretariate and address : Prof. E. Ingelstam, Institute of Optical Research, Stockholm 70.

First presentation of the Stockholm Conference 24th August — 30th August 1959 arranged by the International Commission for Optics (ICO) in connection with its Fifth Meeting, as issued by the ICO Bureau, January 1958. Reference to Professor E. Ingelstam, Stockholm 70.

Modern systems for detecting and evaluating optical radiation.

Progress in this field is due mainly to two trends of development. One is the advent of new receivers, especially for the infrared, and of important electronic aids. The other is the new way of thinking on concepts of energy and information transfer. The entire system must in many respects be considered and treated as consisting of interconnected optical, electronic and other units or sub-systems. It is often designed to achieve improved performance for utilizing weak signals in the presence of a noise background. These topics will form the main theme of the conference, but emphasis will be placed on the broad principles and on methods for their practical application, rather than on the description of specific components.

Bibliographie

Herbert Schober, *Das Sehen*, Bd. 1. 2. Aufl. Fachbuchverlag Leipzig 1957, 367 S., 71 Abb. Preis DM. 28.

Ein Blick auf das Literaturverzeichnis bestätigt, dass an guten und modernen Darstellungen der Ophthalmologischen Optik kaum Mangel ist. Ein weiteres Werk dieser Art findet also eine Daseinsberechtigung nur dann, wenn es sich durch eine besondere Eigenart von den anderen Büchern abhebt. Und wirklich : Der Physiker und Techniker, der für seine Arbeit die Grundtatsachen der physiologischen Optik benötigt, sieht sich in den Sammelwerken über den Gesichtssinn einer ungeheuren Menge von Material gegenüber, das in erster Linie in dem Gedanken an den Mediziner zusammengetragen worden und in seiner Sprache geschrieben ist, die dem Techniker weitgehend unverständlich bleibt. Und umgekehrt, der Mediziner findet kaum ein Buch über die physikalischen Grundlagen der Lichttechnik, das ihn anspricht. In seiner Doppelrolle als Arzt und als Physiker hat der Verfasser die Möglichkeit, die ophthalmologische Optik von beiden Blickpunkten zu betrachten und zu behandeln. So findet der Mediziner alles, was ihn an physikalischen und technischen Tatsachen zu wissen nützlich ist, geschrieben von einem Ophthalmologen. Und der Techniker lernt in gleicher Weise Anatomie und Physiologie des Auges so kennen, wie es seinen Bedürfnissen entspricht. Dieser besondere Vorzug des Werkes wird es zu dem ständigen Ratgeber an beiden Schreibtischen werden lassen. Dort, wo die Einzelheiten Ziel und Rahmen des Buches sprengen würden, gibt das umfangreiche und wohlgeordnete Literaturverzeichnis — 74 Seiten — die nötigen Fingerzeige und Hinweise.

Das Werk kann durch die Kommissionsbuchhandlung Kawe, Berlin-Charlottenburg 2, Bahnhof Zoo, bezogen werden.

Franke.

N° d'éditeur 145. — *Revue d'Optique.* — *Le Directeur* : Emile Perrin.
Commission paritaire des papiers de Presse n° 29754

3978. — Imprimerie JOUVE, 15, rue Racine, Paris. — 9-1958